

Result No.	Score	Query Match Length	DB ID	Description
1	363	16.7	ABP51639	Abp51639 Escherichia coli 9
2	360	16.5	AA40926	AA40926 Wild-type Plasmid p
3	360	16.5	AAWF8898	AAWF8898 Plasmid p
4	359	16.5	AAW80734	AAW80734 Amino aci
5	359	16.5	AAU79467	AAU79467 phoN23097
6	359	16.5	AAE15450	Aae15450 Lac repre
7	359	16.5	AAK40927	Aar40927 Cpg depre
8	359	16.5	ABP51638	Abp51638 Nucleotid
9	310	14.2	ABP43395	ABP43395 E. coli f
10	309	14.2	ABP43374	ABP43374 E. coli g

XX	Disclosure: Page 107-109; 112pp; English.	Qy	6259 GCGGACTGGAGTGCATGTCGGTTTCACAAACCATGCAATTCTGAATGAGGGCATT 6318
XX	The present invention describes Corynebacterium glutamicum transnational regulatory region polynucleotide sequences (I). ABO73173 to ABQ73194 represent the C. glutamicum transcriptional regulatory regions pta, aceB, adh, aldB, PoxB, ldh, amyB, malZ, bgfX, gam, glnX, hisD, pyru, purD, hrcA, hpx, dnaK, ctc, grpE, clpB, and nara, respectively. (I) can be used for producing a vector. (I) can also be used for producing a transformed Corynebacterium sp. host cell, which is useful for producing a biosynthetic product. (I) is useful for producing a Corynebacterium sp. host cell, which is useful for producing amino acid. (I) is useful for regulating and enhancing the production of a variety of products in host cells, including amino acids such as lysine, purine nucleotides such as inosinic acid, and heterologous polypeptides. The present sequence represents an E. coli lacI repressor protein from the present invention	Db	221 GlyAlaPrrSerAlaLekSerGlyTheGlnIntrMetGlnMetLeuAngGlyIle 240
CC		Qy	6319 GTTCCCACTGGATCTGGTGGCCACGATCACAGGGCTGGCCAAATGCGGCATT 6378
CC		Db	241 ValProThrAlaMetLeuAlaAsnAspGlnMetAlaLeuGlyAlaMetArgAlaile 260
CC		Qy	6379 ACCAGTCGGAGTCGGTGGATATCTGGTAGTGGAATGAGGATACGGAA 6438
CC		Db	261 ThrGluSerGlyLeuArgValGlyAlaAspLysSerValValGlyItyraspHrGlu 280
CC		Qy	6439 GACAGCTCACGTTATAATCCCGCGTCAACCACCAACGGAATTTCGCTGGGG 6498
CC		Db	281 AspSerSerCystYleProProSerThrIleProSerLeuGlyLeuGly 300
CC		Qy	6499 CAAACCAAGCGCTGGACGGCTGCTGAACTCTCTCAGGGCOAGGCCGTGAAAGGCATCG 6558
CC		Db	301 GluThrSerValAspArgLeuLeuSerGlnGlyGlnAlaVallysGlyGln 320
CC		Qy	6559 CTGTCGCGCTCTACTGGTAAAGAAACCCCTGCGCCATAGCCAAACGCC 6618
CC		Db	321 LeuLeuProValSerLeuValLysArgLysThrThrLeuAlaProAsnThrGlnThrAla 340
CC		Qy	6619 TCTCCCGCCGGTGGCCGATTCATTATCAGCTGGCACGACAGCTTCGGACTGGAA 6678
CC		Db	341 SerProArgAlaLeuAlaAspSerLeuMetGlnLeuAlaSerArgLeuGlu 360
SQ	Sequence 363 AA:	Qy	6679 AGCGGCCAG 6687
Qy	5599 GTGGGTGATGCAACCGTAACTGTTACGATGTCAGAGTATGCCGTGCTCTTAT 5658	Db	361 SerGlyGln 363
Db	1 ValVaAspValLysProVaThrLeuTyrsPvaAlaGluTyvAlaGlyvA SerTy 20	RESULT 2	
Qy	5659 9AGACCCSTTCCC GGCGTGAATTACATTCCAAACCGCGTGCACACAACCGGGAAAA 5718	ID	AAR40926 standard; protein; 360 AA.
Db	21 GluThrValSerArgValVaAsnGlnAlaLysSerHsvalSerAlaLysThrArgGluLys 40	XX	
Qy	5719 GTGGAAAGCGGCCGATGGCGGAGCTGTAATTACATTCCAAACCGCGTGCACACAACCGGG 5778	AC	AAR40926;
Db	41 valGluAlaAlaMetAlaGluLeuLysThrProAsnArgValAlaGlnGlnLeuAla 60	XX	
Qy	5779 GGCAAACAGCTGCTTGTGATGCCGTTGCCACCTCCAGTCGGCCCTGACGGCGCTCG 5838	DT	25-MAR-2003 (revised)
Db	61 GluLysGlnSerLeuLysLeuLysGlyValAlaLysSerGlnAlaAspGlnLeuGlyAlaSerValValSer 80	DT	09-JAN-2003 (revised)
Qy	5839 CAAATGTCGGCGGATTAATCTCGGCCATCACTGGTGGCTGCTGCTGCTGCTG 5898	XX	22-FEB-1994 (first entry)
Db	81 GluLeuAlaAlaAlaLysLysSerGlnAlaAspGlnLeuGlyAlaSerValValSer 100	XX	
Qy	5899 ATGGTAGAACGAAAGCGCCGTCGAAGCTGTAAGGGCGGTGCACATCTCTCGCGCAA 5958	DE	Wild-type lacI.
Db	101 MetValGluArgSerGlyIvaGluAcysIysAlaAlaValHisAsnLeuAlaGln 120	XX	Bacteria; CpG doublet; vertebrate mammalian; methylation; marker;
Qy	5959 CGCGTCAGTGGGACTGATCATTAACATCCTGCTGGATGACCGAGATCCATGCTGCGTGGAA 6018	KW	reporter; transgene; mutagen; carcinogen; mutation; assay;
Db	121 ArgValSerGlyLeuLeuLysLeuLysSerGlnAlaAspGlnLeuGlyAlaSerValValSer 140	XX	chemical agent.
Qy	6019 GCTGCCCTGCACTAAATGTTCCGGCTTATTCCTGATGTTCTGCTGCTGCTG 6138	OS	Bacteriophage lambda.
Db	141 AlaAlaLysThrAsnValProAlaLeuPhenLeuAspValSerLeuAla 160	XX	WO9317123-A1.
Qy	6079 AGTATTATTTCTCCATGAAAGACGGTAACTGGCGACTTGGCGTGGAGCATCTGGTGCATG 6138	PR	27-FEB-1992; 92US-00842664.
Db	161 SerIleIpheSerHisGluAspGlyIysLysLeuAlaLysLeuAla 180	XX	(OHIS) UNIV OHIO STATE.
Qy	6139 GGTCACTGCAAAATCGCGCTGTTAGCTGCCATTAAAGTCTGCTGCGCGT 6198	PD	02-SEP-1993.
Db	181 GlyHisGlnGlnIleAlaLeuLeuAlaGlyIoleUserSerValSerAlaArgLeuArg 200	XX	Stambrook PJ;
Qy	6199 CTGGCTGCTGGCATATAATCTCACTGCCATCAATTCTGCGGAACGGAA 6258	XX	WPI; 1993-288426/36.
Db	201 LeuAlaGlyT-PhiLysIysLeuThrArgAsnGlnIleGinProLeAlaGluArgGlu 220	XX	N-PSDB; AAQ48615.
Qy		PT	Modified mutagenicity test - comprises assaying transgenic animals for genetic mutations, utilizes native amount of CpG nucleotide(s).
Db		PT	Disclosure; Page 52-53; 79pp; English.
CC		CC	Bacterial genes exhibit a much higher frequency of occurrence of CpG doublet than do vertebrate genes. As a result, a bacterial gene
CC		CC	disclosure; Page 52-53; 79pp; English.

Copyright GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 2, 2005, 19:39:42 ; Search time 159.5 Seconds
(without alignments)
6298.603 Million cell updates/sec

Title: US-10-649-433C-1
Perfect score: 2180
Sequence: 1 gtttgcacgtttatcatcgatcgttgcgcattatgtatcg 6729

Scoring table: Xgapext Ygapext Fgapext Delop

Xgapext	60.0	Ygapext	60.0
Ygapext	60.0	Fgapext	60.0
Fgapext	6.0	Delext	7.0
Delop	6.0		

Searched: 513545 seqs, 7449064 residues

Word size: 1

Total number of hits satisfying chosen parameters: 903960

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_n2p.model -DEV=x10
-Q=/cgn2_1/USPETO_spool_p/US1064433/runat_02082005_165026_1844/app_query.fasta_1.6919
-DB=Issued_Patents_AA -QFMT=fastaN -SUFFIX=Oligo.raii -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=o1190 -TRANS=human40_cdi
-LIST=45 -DOALIGN=C -THR SCORE=quality THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=2000000000
-USER=US1064433 @CGN 1 1 178 @CGN 1 165026 1844 -ICPU=6
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=40 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
2: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
3: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
4: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
5: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
6: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
7: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
8: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
9: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
10: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
11: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
12: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
13: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
14: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
15: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
16: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
17: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
18: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
19: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
20: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
21: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
22: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
23: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
24: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
25: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
26: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
27: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
28: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
29: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
30: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
31: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
32: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
33: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
34: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
35: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
36: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
37: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
38: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
39: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
40: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
41: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
42: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
43: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
44: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
45: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
46: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
47: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
48: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
49: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
50: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
51: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
52: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
53: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
54: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
55: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
56: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
57: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
58: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
59: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
60: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
61: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
62: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
63: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
64: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
65: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
66: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
67: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
68: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
69: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
70: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
71: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
72: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
73: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
74: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
75: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
76: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
77: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
78: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
79: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
80: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
81: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
82: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
83: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
84: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
85: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
86: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
87: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
88: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
89: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
90: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
91: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
92: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
93: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
94: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
95: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
96: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
97: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
98: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
99: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
100: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
101: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
102: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
103: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
104: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
105: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
106: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
107: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
108: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
109: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
110: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
111: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
112: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
113: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
114: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
115: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
116: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
117: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
118: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
119: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
120: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
121: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
122: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
123: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
124: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
125: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
126: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
127: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
128: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
129: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
130: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
131: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
132: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
133: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
134: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
135: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
136: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
137: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
138: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
139: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
140: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
141: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
142: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
143: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
144: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
145: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
146: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
147: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
148: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
149: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
150: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
151: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
152: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
153: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
154: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
155: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
156: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
157: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
158: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
159: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
160: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
161: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
162: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
163: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
164: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
165: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
166: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
167: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
168: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
169: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
170: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
171: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
172: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
173: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
174: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
175: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
176: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
177: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
178: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
179: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
180: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
181: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
182: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
183: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
184: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
185: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
186: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
187: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
188: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
189: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
190: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
191: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
192: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
193: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
194: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
195: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
196: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
197: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
198: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
199: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
200: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
201: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
202: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
203: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
204: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
205: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
206: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
207: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
208: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
209: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
210: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
211: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
212: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
213: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
214: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
215: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
216: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
217: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
218: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
219: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
220: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
221: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
222: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
223: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
224: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
225: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
226: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
227: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
228: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
229: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
230: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
231: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
232: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
233: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
234: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
235: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
236: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
237: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
238: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
239: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
240: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
241: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
242: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
243: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
244: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
245: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
246: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
247: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
248: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
249: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
250: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
251: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
252: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
253: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
254: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
255: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
256: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
257: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
258: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
259: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
260: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
261: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
262: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
263: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
264: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
265: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
266: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
267: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
268: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
269: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
270: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
271: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
272: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
273: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
274: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
275: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
276: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
277: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
278: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
279: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
280: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
281: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
282: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
283: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
284: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
285: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
286: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
287: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
288: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
289: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
290: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
291: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
292: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
293: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
294: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
295: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
296: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
297: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
298: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
299: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
300: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
301: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
302: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
303: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
304: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
305: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
306: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
307: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
308: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
309: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
310: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
311: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
312: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
313: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
314: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
315: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
316: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
317: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
318: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
319: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
320: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
321: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
322: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
323: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
324: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
325: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
326: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
327: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
328: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
329: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
330: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
331: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
332: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
333: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
334: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
335: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
336: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
337: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
338: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
339: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
340: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
341: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
342: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
343: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
344: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
345: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
346: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
347: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
348: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
349: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
350: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
351: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
352: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
353: /cgn2_6/ptoatac/1/iaa/6B/COMB.pep:
354: /cgn2_6/ptoatac/1/iaa/PCNTS/COMB.pep:
355: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
356: /cgn2_6/ptoatac/1/iaa/backfiles/pep:
357: /cgn2_6/ptoatac/1/iaa/5A/COMB.pep:
358: /cgn2_6/ptoatac/1/iaa/5B/COMB.pep:
359: /cgn2_6/ptoatac/1/iaa/6A/COMB.pep:
360: /cgn2_6

SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-01676A-2

Alignment Scores:
Pred. No.: 4.56e-298 Length: 360
Score: 360.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.51% Deletes: 0
DB: 5 Gaps: 0

US-10-649-433C-1 (1-6729) x PCT-US93-01676A-2 (1-360)

Qy 5608 GTGAAACAGTAAACGGTTATCGATGTCAGAGTATGCCGTGTCCTATCAGACCCCT 5667
Db 1 ValLeuProValThrLeuAspPheArgLeuGlyGlnThrSer 300

Qy 5668 TCCGGCTGGTGAACCGGCCAGCTTCCTGCCTAAACGGGAAAGTGGAAAGCG 5727
Db 21 SerArgValValAsnGlnAlaSerHsValSerAlaLysThrArgGluLysValGluAla 40

Qy 5728 GCGATGCCGACCTGTAATTCACTCCCAACCGGTGGCACACAACGTGGGCCAACAG 5787
Db 41 AlaMetAlaGluLeuAsnTyrlProAsnArgValAlaGlnGlnLeuAlaGlyySGln 60

Qy 5788 TCGTTGCTGATGGCCTGCCACCTCCAGCTGGCCCTGCAAGCGCCGTGCCAAATTGTC 5847
Db 61 SerLeuLeuGlyValAlaThrSerSerLeuAlaLeuHsAlaProSerGlnLeuAla 80

Qy 5848 GGGCGATTAATCTGGCCCATCACTGGCTGGCCAGCTGGTGCTGCGATGGTAGAA 5907
Db 81 AlaAlaLysSerArgAlaAspGlnLeuGlyAlaSerValValValSerMetValGlu 100

Qy 5908 CGAAGCGCGTGTGAAGGCTGTAAGCGGGCTGGCACATCTTCCTCCGCAACGGCTCAGT 5967
Db 101 ArgSerGlyValGluLysLysAlaAlaValHsBasilLeuGlyAlaGlyAlaValSer 120

Qy 5968 GGGCTGATCATTAACATCCCGCTGGATGACAGGATGCAATTGGCTGGAGCTGCTGC 6027
Db 121 GlyLeuLeileAsnTyroProLeuAspAspGlnAspAlaAlaAlaCys 140

Qy 6028 ACTAAAGTTCGGCTTATTCCTGATGCTGCTGCAAGAACCCATCACGTTATT 6087
Db 141 ThrAsnValProAlaLeuAspValSerAspGlnThrProLeuAsnSerLele 160

Qy 6088 TTCTCCCCATGAGACGTACGGACTGGCTGGACATCTGGTCCTACGGTCACAG 6147
Db 161 PheSerHisGluAspGlyThrArgLeuGlyValGluHisLeuValAlaLeuGly 180

Qy 6148 CAAATCCGGCTTACGGGCCATTAAGGTTCTGGCTGGCGCTGCGCTGAC 6207
Db 181 GluIleLeuLeuAlaGlyProLeuSerValSerAlaArgLeuArgLeuAlaGly 200

Qy 6208 TGGCATATAATCTCACTGGCAATCAAACCATGCGCATGGCCATAGGGAAAGGGCACTGG 6267
Db 201 TrpHisIleStryLeuLeuArgAsnGlnProLeuAlaGluArgGluGlyAspTrp 220

Qy 6268 AGTGCATGGTCGGTTACGGGCCATTAAGGAAATGCTGAATGAGGCACTGGTCCCCT 6327
Db 221 SerAlaMetSerGlyPheGlnGlnThrMetGlnMetLeuLeuGlyLeuAlaProTrp 240

Qy 6328 GCGATGCTGGTGCACAGATGAGATGGCTGGGGCAATACCGAGTCC 6387
Db 241 AlaMetLeuValAlaAlaAspGlnMetAlaLeuGlyAlaMetArgAlaIleThrGluSer 260

Qy 6388 GGGCTGGCGCTGGGATATGGTGTGGGATATGGGATAAGGAGACAGCTCA 6447
Db 261 GlyLeuIgvaGlyAlaAspIleSerValValGlyTyroAspAspThrGluAspSer 280

Qy 6448 TGTTATATCCGGCTCAAACCATCAAACAGGATTTCCTGGCTGGGGCAAACAGC 6507
Db 281 CysTyrlProProSerThrIleSerIlePheArgLeuGlyGlnThrSer 300

Qy 6508 GTGGACCGCTTGTGCAACTCTCTAGGGCCAGGGCTGAAAGGCCAATCAGCTGTTGCC 6567
Db 301 ValAspArgLeuLeuGlnLeuSerGlyAlaValAlaValGlyAlaGlnLeuLeuPro 320

Qy 6568 GTCTCACTGGTGAAGAAAACCCCTGCCGCCATAGGGAAACCGCTCTCCCGC 6627
Db 321 ValSerIleValLysArgLysThrLeuAlaProAsnThrGlnThrAlaSerProArg 340

Qy 6628 GGTRGGCCGATTAATGAGCTGGCACAGATTCGGACTGGAAAGGGCAG 6687
Db 341 AlaLeuIleAspSerIleMetGlnLeuAlaArgGlnValSerArgLeuGluSerGlyGln 360

RESULT 2
PCT-US93-01676A-4
; Sequence 4, Application PC//TUS9301676A
; GENERAL INFORMATION:
; APPLICANT: Scembrook, Peter J.
; TITLE OF INVENTION: MUTAGENICITY TESTING USING REPORTER
; TITLE OF INVENTION: GENES WITH MODIFIED METHYLATION FREQUENCIES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seven Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01676A
; FILING DATE: 19930226
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/842,644
; FILING DATE: 02-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cooper, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: STAMBOOK 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-62-5197
; TELEX: 248653
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 370 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-01676A-4

Alignment Scores:
Pred. No.: 3.23e-297 Length: 370
Score: 359.00 Matches: 359
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.47% Indels: 0
DB: 5 Gaps: 0

US-10-649-433C-1 (1-6729) x PCT-US93-01676A-4 (1-370)

Qy 5611 AAACAGTAGTAACTGGTTACGATGTCGCGAGACTATCGACCCSTTTC 5670
Db 2 LysProValThrLeuLeuAspValAlaGluItyrAlaGlyValSerTyrginThrValSer 21

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 2, 2005, 20:07:07 ; Search time 801 Seconds

(without alignments)
6545.998 Million cell updates/sec

Title: US-10-649-433C-1

Perfect score: 2180

Sequence: 1 gtttgacagcttcatatcgaa.....agttagcgccaaattgtatctg 6729

Scoring table: OLIGO

Xgapext 60.0 , Xgapext 60.0

Ygapext 60.0 , Ygapext 60.0

Fgapext 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1745140 seqs., 389608008 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3345576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p model -DEV=x1D
-Q=/cpn2_1/usfPRO_apool_P/US10649433/runat_02082005_165027_1863/app_query.fasta_1.6919
-DB=Published_Applications_AA -QFORMAT=tastan -SUFFIX=Oligo_rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIXX=oligo
-TRANS:human140_cdi -LIST=45 -DOALIGN=200 THR SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTWT=0 -PRT=NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=00000000 -USER=0000000000 @CGN 1. 964 <runat 02082005_165027_1863
-NCPU=6 -ICPU=3 NO_NMAP -LARGEQUERY NEG_SCORES=0 -WAIT -DSPBLOCK=100
-DLONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=60 -XGAPOF=60 -XGAPEXT=60
-RGAPOP=6 -RGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7
```

Database :

Published Applications AA: *

1: /cgn2_6/ptodata/1/pubcaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubcaa/FCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubcaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubcaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubcaa/US07_PUB.pep:*

6: /cgn2_6/ptodata/1/pubcaa/BCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubcaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubcaa/US09_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubcaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/picodata/1/pubcaa/US09C_PUBCOMB..pep:*

11: /cgn2_6/picodata/1/pubcaa/US10A_PUBCOMB..pep:*

12: /cgn2_6/picodata/1/pubcaa/US10B_PUBCOMB..pep:*

13: /cgn2_6/picodata/1/pubcaa/US10C_PUBCOMB..pep:*

14: /cgn2_6/picodata/1/pubcaa/US10B_PUBCOMB..pep:*

15: /cgn2_6/picodata/1/pubcaa/US10C_PUBCOMB..pep:*

16: /cgn2_6/picodata/1/pubcaa/US10D_PUBCOMB..pep:*

17: /cgn2_6/picodata/1/pubcaa/US10E_PUBCOMB..pep:*

18: /cgn2_6/picodata/1/pubcaa/US10_F_PUBCOMB..pep:*

19: /cgn2_6/picodata/1/pubcaa/US11A_PUBCOMB..pep:*

20: /cgn2_6/picodata/1/pubcaa/US11_NEW_PUB..pep:*

21: /cgn2_6/picodata/1/pubcaa/US60_NEW_PUB..pep:*

22: /cgn2_6/picodata/1/pubcaa/US60_PUBCOMB..pep:*

ALIGNMENTS

RESULT 1 US-09-987-763-29
/ Sequence 29, Application US/09987763
/ Publication No. US003001755A1
/ GENERAL INFORMATION:
/ APPLICANT: Rayapati, P. John
/ ATTORNEY: Crafton, Corey M.
/ TITLE OF INVENTION: Nucleotide Sequences for Transcriptional Regulation in Corynebacterium glutamicum
/ TITLE OF INVENTION: 1533.1.1940002
/ FILE REFERENCE: CURRENT APPLICATION NUMBER: US/09/987-763
/ CURRENT FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: 60/248,219
/ PRIOR FILING DATE: 2000-11-15
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: Patentnet version 3.0
/ SEQ ID NO 29
/ LENGTH: 363
/ TYPE: PRT
/ ORGANISM: Escherichia coli

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	FEATURES	Source	Location/Qualifiers
1	3855	57.3	4806	12 AF050464	AF050464 Expression	1..4806		/organism="Expression vector pKIL-HIS3"
2	3804	56.5	5013	6 AR24513	AR24513 Sequence	250..774	CDS	/mol_type="genomic DNA"
3	3753	55.8	4176	6 A13038	A13038 Artificial	/note="specific host="Escherichia coli"		/note="CcdB fusion protein"
4	3753	55.8	4176	12 SYNTRCP9A	M22744 Cloning vec	28..1002	gene	/db_xref="taxon:74458"
5	3753	55.8	4239	6 A29289	A29289 PSEC-Bp1	/gene="ccdB"		/codon_start=1
6	3753	55.8	5271	12 CVD089936	U89936 Cloning vec	/transl_table=11		/product="CcdB protein"
7	3752	55.8	4476	6 AR102989	AR102989 Sequence			/db_xref="GI:3892176"
8	3752	55.8	4476	6 AR172793	AR172793 Sequence			/translation="MSRRRRQAWFSSSSSLIEGSVDLAVLQRDRWWNPGBTQLNRL"
9	3703	55.0	4214	6 BD273391	BD273391 P450/acet			
10	3703	55.0	4214	6 AX027488	AX027488 Sequence			
11	3702	55.0	4176	12 XXU13872	U13872 pTRC99a Clo			
12	3702	55.0	4203	6 AX576404	AX576404 Sequence			
13	3702	55.0	4204	6 AX576405	AX576405 Sequence			
14	3702	55.0	4205	6 AX576406	AX576406 Sequence			
15	3702	55.0	4216	6 AX576407	AX576407 Sequence			
16	3702	55.0	4217	6 AX576408	AX576408 Sequence			
17	3702	55.0	4218	6 AX576409	AX576409 Sequence			
18	3384	50.3	4232	12 SYNBETACQ	L24367 Cloning vec			
19	3384	50.3	4735	12 SYNBETACQ	L24367 Cloning vec			

ALIGNMENTS

RESULT 1
AF050464
LOCUS AF050464 Expression vector pKIL-HIS3, complete sequence.
DEFINITION Expression vector pKIL-HIS3
ACCESSION AF050464
VERSION AF050464.1 GT:3892175

SOURCE Expression vector pKIL-HIS3
ORGANISM Expression vector pKIL-HIS3
OTHER SEQUENCES; ARTIFICIAL SEQUENCES; VECTORS.
REFERENCE 1 (bases 1 to 4806)
AUTHORS Van Reeth,T., Dreze,P.L., Szipiro,J., Gabant,P.
TITLE Positive selection vectors to generate fused genes for the expression of his-tagged proteins
JOURNAL Biotechniques 25 (5), 898-904 (1998)
MEDLINE 99039034
PUBMED 9821593
REFERENCE 2 (bases 1 to 4806)
AUTHORS Van Reeth,T., Gabant,P., Dreze,P.L., Szipiro,J. and Szpirer,C.
TITLE Direct Submission
JOURNAL Universite Libre de Bruxelles, 67, rue des Chevaux, Rhode St.
GENE 1640, Belgium
GENE 1640, Belgium
SUBMITTED (24-FEB-1998) Departement de Biologie Moleculaire,
UNIVERSITE LIBRE DE BRUXELLES, 67, RUE DES CHEVAUX, RHODE ST.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3855	57.3	4806	12 AF050464	AF050464 Expression
2	3804	56.5	5013	6 AR24513	AR24513 Sequence
3	3753	55.8	4176	6 A13038	A13038 Artificial
4	3753	55.8	4176	12 SYNTRCP9A	M22744 Cloning vec
5	3753	55.8	4239	6 A29289	A29289 PSEC-Bp1
6	3753	55.8	5271	12 CVD089936	U89936 Cloning vec
7	3752	55.8	4476	6 AR102989	AR102989 Sequence
8	3752	55.8	4476	6 AR172793	AR172793 Sequence
9	3703	55.0	4214	6 BD273391	BD273391 P450/acet
10	3703	55.0	4214	6 AX027488	AX027488 Sequence
11	3702	55.0	4176	12 XXU13872	U13872 pTRC99a Clo
12	3702	55.0	4203	6 AX576404	AX576404 Sequence
13	3702	55.0	4204	6 AX576405	AX576405 Sequence
14	3702	55.0	4205	6 AX576406	AX576406 Sequence
15	3702	55.0	4216	6 AX576407	AX576407 Sequence
16	3702	55.0	4217	6 AX576408	AX576408 Sequence
17	3702	55.0	4218	6 AX576409	AX576409 Sequence
18	3384	50.3	4232	12 SYNBETACQ	L24367 Cloning vec
19	3384	50.3	4735	12 SYNBETACQ	L24367 Cloning vec

missc_feature									
terminator									
gene									
CDS									
/gene="ccdB"									
/locus_tag="NruI_his-tagged_protein"									
253..259									
/note="unique NruI (blunt) cloning site that can be used for generation of C-terminal His-tagged protein"									
876..1102									
/gene="ccdB"									
1313..2173									
/locus_tag="ampR"									
1313..2173									
/gene="ampR"									
/codon_start=1									
/transT_table=11									
/product="beta-lactamase"									
/protein_id="AAC9291_1"									
/db_xref="GI:38922177"									
/translation="MSIICHRVLLPPIFAACFLCPVPAHPEPTLVKVKDAEOLGARVGV IEELDLSKILLESPPRPERPPMSTEVKVLCCAVLNSRDAGQEQLGRRHYSQNDLVE YSPVTEKLTDGNTVRECSAATIMSDNTAAANLLTTGGPKELTLASRQQLIDMMEADKVAGPL DRQPELNEAIPINDERDTMPVAMATTLRKLTGEULLASRQQLIDMMEADKVAGPL LRSRDLPAQMFIALDKSGAERGSRGIIIAALGFDGKPSRIVVYITTSQATMDERNRQIA EIGASLRLKHW"									
2319..2992									
3645..4604									
/gene="lacIQ"									
3645..4604									
/gene="lacIQ"									
/function="regulator of lac promoter"									
/codon_start=1									
/transT_table=11									
/product="LaciQ"									
/protein_id="AAC78292_1"									
/db_xref="GI:3892178"									
/translation="MAENYIINRVAQOLAKGKOSLLIGVATSSLALHAPSQIVAAIKS RAHQLGAVSYVSVERSEVACKAAYVNLIAQRVSHELIINYPLDDODAIAVEAATCNY PAFLDLDSDQTIPINSIIFSHEDOTRGLHLQHGAATTCAGTTTCAAGATTAATCAGAC HKYLTRNDIQOPTAIREGGWSAMSFGQTMOMLNGEVTPVAMILVANDOMALGAMRATIE SGURVGAIISVYGYDDPESCYIPPSTETIKODFRUIGQTSYDRLQLSQCAVKNGQ LILPVLVRKRKTILAPNTQATASPRALADSMLQARQVSRLBSGQ"									
ORIGIN									
Query Match	57.3%	Score 3855;	DB 12;	Length 4806;					
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;				
Matches 3855;	Conservative 0;	Missmatches 0;							
Qy	2875	AGCTTGCGCTGTTGGCGGTAGAGGAAATTGTTGATCAGATTAATCAGAC				2934			
Db	789	AGCTTGCGCTGTTGGCGGTAGAGGAAATTGTTGATCAGATTAATCAGAC				848			
Qy	2935	GCAGAAGCGCTCTGTATAAACAGAAATTGCTGCCAGTAGTGCGCTGGCTCCACTG				2994			
Db	849	GCAGAAGCGCTCTGTATAAACAGAAATTGCTGCCAGTAGTGCGCTGGCTCCACTG				908			
Qy	2995	ACCCATGCCAACCTGAACTGAGGTGAACGCCATGTTGCTGGCTGGCTCC				3054			
Db	909	ACCCATGCCAACCTGAACTGAGGTGAACGCCATGTTGCTGGCTGGCTCC				968			
Qy	3055	ATGCCAGAGTAGGGAACTGGCAGGATCAAATAAACGAAAGGCTCAGTCGAAAGACTGG				3114			
Db	969	ATGCCAGAGTAGGGAACTGGCAGGATCAAATAAACGAAAGGCTCAGTCGAAAGACTGG				1028			
Qy	3115	GCCTTCGTTTATCTGTTGCTGGCGTAACCTCTCTGAGTAGACAATCCTCCCG				3174			
Db	1029	GCCTTCGTTTATCTGTTGCTGGCGTAACCTCTCTGAGTAGACAATCCTCCCG				1088			
Qy	3175	GGAGCGGATTGAACTTGGCAAGAACGCCGCGGAGGTGGCGGAGGACCCGC				3234			
Db	1089	GGAGCGGATTGAACTTGGCAAGAACGCCGCGGAGGTGGCGGAGGACCCGC				1148			
Qv	3235	TAACCTGGCCAGGCCATCAAAATTAGGAGAAGGCCATCCGAGGATGGCTT				3294			

Run on:	August 5, 2005, 18:49:20 ; Search time 3194 Seconds (without alignments) 12471.450 Million cell updates/sec	
Title:	US-10-649-433C-1	
Perfect score:	6729	
Sequence:	1 gtttgcacgttcatatcgaaaaaaaaatgtatctg 6729	
Scoring table:	OLIGO_NUC	
	Gapop 60.0 , Gapext 60.0	
Searched:	4390206 seqs, 2959870667 residues	
Word size :	0	
Total number of hits satisfying chosen parameters:	8780412	
Minimum DB seq length:	0	
Maximum DB seq length:	2000000000	
Adg46817 Donor Plasmid		
Aaa5260 Nucleotid		
Aaa62632 Vector pE		
Aan90709 Sequence		
Adi136476 Plasmid		
Abq74925 Avidity I		
Abq74925 Avidity I		
Abq74926 Avidity I		
Abq74927 Avidity I		
Abq74928 Avidity I		
Abq74929 Avidity I		
Abq74930 Avidity I		
Adg46818 PEM DNA		
Adg47070 Mouse OBQ		
Aal55822 Phi-29 expression		
Aal55823 Phi-29 expression		
Ado37103 Vector pE		
Ado77625 pBAD/GII		
Aaf31389 Expression		
Aaq05397 Secretion		
Aal48883 B subtilis		
Aal11227 Pantoea		
Adl92656 Expression		
Adt42002 VAP related		
Aaa07775 DNA sequence		
Abz76630 HIF-1 alp		

ALIGNMENTS

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	Start			
1	3855	57.3	5027	2	AAX78865	Human tis	Aax78865 Human tis
2	3855	57.3	5069	2	AAX78871	Human tis	Aax78871 Human tis
3	3855	57.3	5903	8	BAB23939	Nucleotid	Ab23939 PTRCHisB
4	3804	56.5	4816	4	AAD07799	Mouse Adi	Ad07799 PTRCHisB
5	3804	56.5	4816	12	ADF72546	Mouse Adi	Adf72546 PTRCHisB
6	3804	56.5	4816	12	ADG14708	Globular	Adg14708 PTRCHisB
7	3804	56.5	5013	9	ACA63355	E. coli D	Aca63355 E. coli D
8	3804	56.5	5013	10	AAD59423	E. coli Pu	Aad59423 E. coli Pu
9	3804	56.5	5013	10	ADE6127	E. coli P	Ade6127 E. coli P
10	3804	56.5	5104	4	AAD07798	PTRCHisB	Ad07798 PTRCHisB
11	3804	56.5	5104	12	ADF72545	Mouse Adi	Adf72545 PTRCHisB
12	3753	55.8	5069	2	AAX78874	Human tis	Aax78874 Human tis
13	3753	55.8	5099	2	AAX78896	Human tis	Aax78896 Human tis
14	3753	55.8	5132	2	AAX78883	Human tis	Aax78883 Human tis
15	3753	55.8	5324	2	AAX78890	Human tis	Aax78890 Human tis
16	3753	55.8	5342	2	AAX78893	Human tis	Aax78893 Human tis
c 17	3753	55.8	5926	2	AAV32978	Tn7 donor	Aav32978 Tn7 donor
c 18	3753	55.8	5926	2	AAV32977	Tn7 donor	Aav32977 Tn7 donor
c 19	3753	55.8	5926	6	AAD43060	PEM delta	Aad43060 PEM delta
c 20	3753	55.8	5926	6	AAD43059	Transpo	Aad43059 Transpo

PreP. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

/product = "TF clone NuV120"

XX	PN	WO9932143-A1.
XX	XX	
XX	PD	01-JUL-1999.
XX	PR	23-DEC-1998;
XX	PR	23-DEC-1997;
XX	PR	97US-00996744.
XX	(NUVA-)	NUVAS LLC.
XX	PI	Houston LL, Dickinson CD;
XX	DR	WPI; 1999-405116/34.
XX	DR	P-PSDB; AAY25403.
XX	PT	New thrombogenic polypeptides used to, e.g. obliterate vascular malformations.
XX	PT	Claim 30; Page 68-70; 97pp; English.
XX	CC	This invention describes novel thrombogenic polypeptides which comprise a thrombogenic substructure and a context-dependent entity which recognizes a desired biologically susceptible sites. e.g. tumour vascular endothelium
CC	CC	
CC	CC	
CC	CC	

CC A novel context-dependent functional entity comprises a substructure with
 CC thrombogenic potential and one or more context-enhancing substructures
 CC having the ability to recognize desired biologically susceptible sites,
 CC where the entity imparts thrombogenic activity when positioned in the
 CC function-forming-context at the biologically susceptible sites, and the
 CC entity has no thrombogenic activity absent a function-forming-context at
 CC the biologically susceptible sites. The context-dependent functional
 CC entities impart thrombogenic activity only at biologically susceptible
 CC sites. They can be used to obliterate vascular malformations or to
 CC selectively thrombose the vasculature of solid tumours. This sequence
 CC encodes the human tissue factor protein Nuv120 which is used in the
 CC method of the invention

xx Sequence 5027 BP; 1287 A; 1221 C; 1321 G; 1198 T; 0 U; 0 Other;

Query Match	57.3%	Score 3855;	DB 2;	Length 5027;	
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Gaps 0;	
Matches 3855; Conservate	0;	Indels 0;			
Qy 2875 AGCTTGCCTGTTGGGGATGAGAAGATTTCAGCCTGATACAGATAAATCAGAAC 2934	Db 1173 AGCTTGCCTGTTGGGGATGAGAAGATTTCAGCCTGATACAGATAAATCAGAAC 1232	Qy 2935 GAGAAAGCGGTCTGATAAACAGAAATTGCTCTGGCGTAGCGCGATGGTAGTCCCACCTG 2994	Db 1233 GCAAGAAGGGCTGCTGATAAACAGAAATTGCTCTGGCGTAGCGCGATGGTAGTCCCACCTG 1292	Qy 2995 ACCCATGCGGAACACTGAAAGTGAAACGCCGTAGCGCGATGGTAGTCCCACCTG 3054	Db 1293 ACCCATGCGGAACACTGAAAGTGAAACGCCGTAGCGCGATGGTAGTCCCACCTG 1352
Qy 3005 ATGCGAGAATGAGGGAACTGCCAGGCAATTAACAGAAGGGCTCACTGCGAAAAGACTGG 3114	Db 1353 ATGCGAGAATGAGGGAACTGCCAGGCAATTAACAGAAGGGCTCACTGCGAAAAGACTGG 1412	Qy 3115 GCCTTGGTTATCGTTGTTGCGGTGAGCTCTCTGAGTAGGACAATTCGCGG 3174	Db 1413 GCCTTGGTTATCGTTGTTGCGGTGAGCTCTCTGAGTAGGACAATTCGCGG 1472	Qy 3175 GAGAGGAAATTGAAACTGTTGCGGTGAGCTCTCTGAGTAGGACAATTCGCGG 3234	Db 1473 GAGAGGAAATTGAAACTGTTGCGGTGAGCTCTCTGAGTAGGACAATTCGCGG 1532
Qy 3235 TAAACTCGAGGCAATCAAATTAGCGAGCAAGGCCATCTGACGGATGCCCTTTGGTT 3294	Db 1533 TAAACTCGAGGCAATCAAATTAGCGAGCAAGGCCATCTGACGGATGCCCTTTGGTT 1592	Qy 3295 CTACAAACTCTTTGTTATTTCTAAATACATTCAAATATGTAATCGGTCTCATGAGAC 3354	Db 1593 CTACAAACTCTTTGTTATTTCTAAATACATTCAAATATGTAATCGGTCTCATGAGAC 1652	Qy 3355 ANTAACCGTATAATCTCCATATAATTGAAAGGAAGGTGATTGATTCACATT 3414	Db 1653 ANTAACCGTATAATCTCCATATAATTGAAAGGAAGGTGATTGATTCACATT 1712
Qy 3415 TCCGTGTCGCCCTTATCCCTTTCGGCATTTGCTCTGTTGCTCACCCAG 3474	Db 1713 TCGTGTGCCCTTATCCCTTTCGGCATTTGCTCTGTTGCTCACCCAG 1772	Qy 3475 AACAGCTGCTGAAAGTAAAGATGCTGAAAGATCACTGTTGCTCACCG 3534	Db 1773 AACAGCTGCTGAAAGTAAAGATGCTGAAAGATCACTGTTGCTCACCG 1832	Qy 3595 TGATGAGGACATTAAAGTTGCTGTTGAGCTGAGTGGGGTTACATCG 3654	Db 1893 TGATGAGGACATTAAAGTTGCTGTTGAGCTGAGTGGGGTTACATCG 1952
Qy 3655 AAGAGCAACTCGCTGCGCAPACACTTCTCAGAATGACTTGGTGAFACTCACAG 3714	Db 3033 AAGAGCAACTCGCTGCGCAPACACTTCTCAGAATGACTTGGTGAFACTCACAG 3092				

Db 1953 AAGAGCAACTCGCTGCGCATACACTTCTCAGAATGACTTGGTGAFACTCACAG 2012	Qy 3715 TCAAGAAAGCATCTTACGGATGGATGAGTAAAGTGGAGATTGAGAAGCTTCCAA 3774	Db 2013 TCAAGAAAGCATCTTACGGATGGATGAGTAAAGTGGAGATTGAGAAGCTTCCAA 2072	Qy 3775 CCATGAGTGAACACTGCGCCAACTACTCTGACAACTGCAACGATGGAAAGGAGC 3834	Db 2073 CCATGAGTGAACACTGCSCCTGATCGTGTGGAAACCGG 2132	
Qy 3894 TAACCCCTTTTGGACAAACATGGGGATCATGTAACACTCGCTGACGGAC 3894	Qy 2133 TAACCCCTTTTGGACAAACATGGGGATCATGTAACACTCGCTGACGGAC 2192				
Qy 3895 AGCTGATGAGCCATACAAAGCAGGAGCTGACCCAGTGGCTGACAACTGGCAA 3954	Qy 2193 AGCTGATGAGCCATACAAAGCAGGAGCTGACCCAGTGGCTGACAACTGGCAA 2252				
Db 2253 CAACGTTGCGCAAACTTAACTGGGAACTACTACTCTAGCTCCGGAAACAAATAA 4014	Qy 2313 TAGACTGGATGGGGATAAAGTGGAGACCAACTTCTCGGGATCATCTGGCAG 2312				
Qy 4015 TAGACTGGATGGGGATAAAGTGGAGACCAACTTCTCGGGATCATCTGGCAG 4074	Db 2373 GCTGGTTTATGCTGATAAACTGGGGCTGCTGGGTGCTGCTGGCTTCCGGCTG 2372				
Qy 4075 GCTGGTTTATGCTGATAAACTGGGGCTGCTGGGTGCTGCTGGCTTCCGGCTG 4134	Db 2433 CACTGGGGCAAGATGATAAGCCCTCCGGTATCGTATTATCTACAGCGGGATCTGG 4194				
Qy 4195 CAACTATGGTGAACGAAATTAGACAGATGGCTGAGTAACTGCTACTGTTACAGCG 2492	Db 2433 CACTGGGGCAAGATGATAAGCCCTCCGGTATCGTATTATCTACAGCGGGATCTGG 2492				
Qy 2493 CAACTATGGTGAACGAAATTAGACAGATGGCTGAGTAACTGCTACTGTTACAGCG 2552	Db 2493 CAACTATGGTGAACGAAATTAGACAGATGGCTGAGTAACTGCTACTGTTACAGCG 2552				
Qy 2553 GCTAACTGTCAGACCAAGTTACTGTTACATGTTGATAAACTCTCTTACAGCTTAAAC 4314	Db 2613 ATTAAAGGAACTGAGTGGAGATCTCTGTTGATAATCTGTTGATAATCTGTTGATA 2672				
Qy 2672 ATTAAAGGAACTGAGTGGAGATCTCTGTTGATAATCTGTTGATAATCTGTTGATA 2672	Db 2673 GTGAGTTTGTGTTCACTGACCGTGAACCCGTGAACTGCTGTTGAGTAAAGATCA 2732				
Qy 2732 GTGAGTTTGTGTTCACTGACCGTGAACCCGTGAACTGCTGTTGAGTAAAGATCA 2732	Db 2733 ATCCCTTTCTGCGTAACTCCGCTTCTGCGTGTAACTCCGCTTCTGCGTGTAA 2792				
Qy 2792 ATCCCTTTCTGCGTAACTCCGCTTCTGCGTGTAACTCCGCTTCTGCGTGTAA 2792	Db 2793 TGGTTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2852				
Qy 2852 TGGTTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2852	Db 2853 GAGCGGAGATAACCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2912				
Qy 2912 GAGCGGAGATAACCAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2912	Db 2913 ACTCTGTAAGCACCCTCATACCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGG 2972				
Qy 2972 ACTCTGTAAGCACCCTCATACCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGG 2972	Db 2973 GTGGGCTATAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3032				
Qy 3032 GTGGGCTATAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3032					

ALIGNMENTS

RESULT 1
US-03-183-188B-5
; Sequence 5, Application US/09183188B
; Patent No. 6491905
; GENERAL INFORMATION:
; APPLICANT: Sorbher, Eric J.
; APPLICANT: Parker, William B.
; APPLICANT: Waud, William
; APPLICANT: Gadi, Vijaykrishna K.
; TITLE OF INVENTION: RECOMBINANT E. COLI FOR DELIVERY OF PNP TO TUMOR CELLS
; FILE REFERENCE: UAB-12405/22
; CURRENT APPLICATION NUMBER: US/09/183,188B
; PRIORITY NUMBER: 1996-10-30
; PRIOR FILING DATE: 1996-08-23
; PRIOR APPLICATION NUMBER: 08/122,321
; PRIOR FILING DATE: 1993-09-14
; PRIOR APPLICATION NUMBER: 60/064,676
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 5013
; TYPE: DNA
; ORGANISM: Escherichia coli
; US-09-183-188B-5

Query Match	Score	DB 4;	Length	5013;
Best Local Similarity	100.0%;	Pred. No.	0;	
Matches 3854;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;

Qy 2875 AGCTTGGCTTGTGGGAGTAGAGAAGATTTGCGCTGATACTGAGTAAATCGAGAC 2934
Db 1159 AGCTTGGCTTGTGGGAGTAGAGAAGATTTGCGCTGATACTGAGTAAATCGAGAC 1218
Qy 2935 GCAGAAGCGCTTCGATTAACAGATACTGCGCTGCGGAGTAGCGCTGGTCCACCTG 2994
Db 1219 GCAGAAGCGCTTCGATTAACAGATACTGCGCTGCGGAGTAGCGCTGGTCCACCTG 1278
Qy 2995 ACCCATGCCGAACCTGAGGTGAACGCCGTAGGCCGTAGTGTTGGGTCTCCCC 3054
Db 1279 ACCCATGCCGAACCTGAGGTGAACGCCGTAGGCCGTAGTGTTGGGTCTCCCC 1338
Qy 3055 ATGGAGACTGGGAACCTGCCAGCCTCAATAAACAGAAAGGCTCAGTCGAAAGCTGG 3114
Db 1339 ATGGAGACTGGGAACCTGCCAGCCTCAATAAACAGAAAGGCTCAGTCGAAAGCTGG 1398
Qy 3115 GCCATTGCTTATCTGTGTGTTGCGTGAACGCTCTCTGTGACAATTCGCCG 3174
Db 1399 GCCATTGCTTATCTGTGTTGCGTGAACGCTCTCTGTGACAATTCGCCG 1458

Qy	3175	GGAGCGGAAATTGAAACCTTGGCAAGCCAAACGGCCCCGGAGGGTGGCGGCCAGGACGCCA	3234	Qy	4255	GGTAACCTGTGAGACCGAGTTACTCATATACTTTAGATTGATTAACCTTCATTTT	4314
Db	1459	GGAGCGGAAATTGAAACCTTGGCAAGCCAAACGGCCCCGGAGGGTGGCGGCCAGGACGCCA	1518	Db	2539	GGTAACCTGTGAGACCGAGTTACTCATATACTTTAGATTGATTAACCTTCATTTT	2598
Qy	3235	TAAACTGCAAGCATAAATTAGCGAAGGCCATCTGACCGATGGCCTTTGGTT	3294	Qy	4315	AATTAAAAGGATCTAGTGGAGAACCTTGTGAAATCTCATGACCCAAATCCCTTAAC	4374
Db	1519	TAAACTGCAAGCATAAATTAGCGAAGGCCATCTGACCGATGGCCTTTGGTT	1578	Db	2599	AATTAAAAGGATCTAGTGGAGAACCTTGTGAAATCTCATGACCCAAATCCCTTAAC	2658
Qy	3295	CTACAACCTCTTTTGTATTATTCTAAATCATCAAAATATGATTCGGCTCATGAGAC	3354	Qy	4375	GTGAGTTCTCACTGAGCTAGACCCGTAGAACCTGAAAGATCAAGGATCTCTTGAG	4434
Db	1579	CTACAACCTCTTTTGTATTATTCTAAATCATCAAAATATGATTCGGCTCATGAGAC	1638	Db	2659	GTGAGTTCTCACTGAGCTAGACCCGTAGAACCTGAAAGATCAAGGATCTCTTGAG	2718
Qy	3355	AATAACCCGTATAAATGCTGATAAATTGAAAGGAGAGTGTGATTCACATT	3414	Qy	4435	ATCCATTCTCTGCTGGTAAATCTGTGCTGGTAAACACCGCTACAGCGGG	4494
Db	1639	AATAACCCGTATAAATGCTGATAAATTGAAAGGAGAGTGTGATTCACATT	1698	Db	2719	ATCCATTCTCTGCTGGTAAATCTGTGCTGGTACACAAACACCGCTACAGCGGG	2778
Qy	3415	TCCGTGTCGCCCTTATTCCCTTTTGGGATTTCGCTCTGTCACCCAG	3474	Qy	4495	TGGTTGTTGCCGATCAAGCTTACCAACTCTTTCGGAGGTAACTGGCTTCAGCA	4554
Db	1699	TCCGTGTCGCCCTTATTCCCTTTTGGGATTTCGCTCTGTCACCCAG	1758	Db	2779	TGGTTGTTGCCGATCAAGCTTACCAACTCTTTCGGAGGTAACTGGCTTCAGCA	2838
Qy	3475	AAACGCTGGAAAGTAAAGATGCTGAAGATCAGTGGGGTTACATCG	3534	Qy	4555	GAGCGAGAGTACCAAAATACTGCTCTCTAGTGTGGCTAGTTAGCCACACTCAGA	4614
Db	1759	AAACGCTGGAAAGTAAAGATGCTGAAGATCAGTGGGGTTACATCG	1818	Db	2839	GAGCGAGAGTACCAAAATACTGCTCTCTAGTGTGGCTAGTTAGCCACACTCAGA	2898
Qy	3535	AACTGGATCTCAACAGCGGTAAAGATCTGAGCTTTCGCCCCGAAAGAACGTTTCCAA	3594	Qy	4615	ACTCTTAGACCCCTACATACCTGTGCTGTGTAATCCGTTAACAGTGTGCTGCCA	4674
Db	1819	AACTGGATCTCAACAGGGTAGATCTGAGATTTCGCCCCGAAAGAACGTTTCCAA	1878	Db	2899	ACTCTTAGACCCCTACATACCTGTGCTGTGTAATCCGTTAACAGTGTGCTGCCA	2958
Qy	3595	TGATGAGCACTTTAAAGTCGCTATGTCGGCGGTATATTACCCSTGTTACGCGGGC	3654	Qy	4675	GTGGCGATAACTGTRCTCTACCGGGTTGGACTCTACGACGATAGTTACCCGATAAGGC	4734
Db	1879	TGATGAGCACTTTAAAGTCGCTATGTCGGCGGTATATTACCCSTGTTACGCGGGC	1938	Db	2959	GTGGCGATAACTGTRCTCTACCGGGTTGGACTCTACGACGATAGTTACCCGATAAGGC	3018
Qy	3655	AGAGGAAACTCTGTGGCATACATCTCAGATGACTCTGGTGAATCACAG	3714	Qy	4735	AGCGGTCGGGGTGAACGGGGGTTCGTCGACACGCCACCTTGGCGAACGACTTACA	4794
Db	1939	AGAGGAACTCTGTGGCATACATCTCAGATGACTCTGGTGAATCACAG	1998	Db	3019	AGCGGTCGGGGTGAACGGGGGTTCGTCGACACGCCACCTTGGCGAACGACTTACA	3078
Qy	3715	TCACTGAAAGCATCTACGGATGGCATGACGATGAACTTACGATGAACTAA	3774	Qy	4795	CCTGAACTGAGATACTCTACGGATGGCATGACGATGAACTTACGATGAACTAA	4854
Db	1999	TCACTGAAAGCATCTACGGATGGCATGACGATGAACTTACGATGAACTAA	2058	Db	3079	CCTGAACTGAGATACTCTACGGATGGCATGACGATGAACTTACGATGAACTAA	3138
Qy	3775	CCATGAGCTGATAAACACTCGGGCAACCTACTCTGACAACGATCTGGGACGGCC	3834	Qy	4855	AGCGGACAGGTATGGTAAAGGGCTGGGAGGCTGGGAGGCTGGGAGGCTTC	4914
Db	2059	CCATGAGCTGATAAACACTCGGGCAACCTACTCTGACAACGATCTGGGACGGCC	2118	Db	3139	AGCGGACAGGTATGGTAAAGGGCTGGGAGGCTGGGAGGCTTC	3198
Qy	3835	TAACCGTTTTTGCAAGAACATGGGGATCATGTAACTGCTTGTGAACTGG	3894	Qy	4915	CGGGGGAAAAGGCCCTGGTATTTTAACTGCTTCTGGTCTGGCTTC	4974
Db	2119	TAACCGTTTTTGCAAGAACATGGGGATCATGTAACTGCTTGTGAACTGG	2178	Db	3199	CGGGGGAAAAGGCCCTGGTATTTTAACTGCTTCTGGCTTC	3258
Qy	3895	ACGTGATGAAAGCATACAAACGCGGTGACCCAGTGGCTGAACTGGCAA	3954	Qy	4975	GTGCAATTGTTGATCTGCTGTCAGGGGGCGAGCTATGAAACGCCCGAACGGGG	5034
Db	2179	ACGTGATGAAAGCATACAAACGCGGTGACCCAGTGGCTGAACTGGCAA	2238	Db	3259	GTGCAATTGTTGATCTGCTGTCAGGGGGCGAGCTATGAAACGCCCGAACGGGG	3318
Qy	3955	CAACGTTGCGAAACATTAATCTGGGAAACTACTTACTCTGGGAACTACATTAA	4014	Qy	5035	CCTTTTACGGTTCTGGCCTTTGGCTCTCATGTTCTGGCTTCTGGCTTAT	5094
Db	2239	CAACGTTGCGAAACATTAATCTGGGAAACTACTTACTCTGGGAACTACATTAA	2298	Db	3319	CCTTTTACGGTTCTGGCCTTTGGCTCTCATGTTCTGGCTTAT	3378
Qy	4015	TGACTGTTGAGGGGATAAAGTCGAGGACCACTCTGGCTCCGGTACAGGG	4074	Qy	5095	CCCCGATTCTGGTATAACCGGTATAACCGCTTGTGAGCTGATACCCGCTCGCGCA	5154
Db	2299	TGACTGTTGAGGGGATAAAGTCGAGGACCACTCTGGCTCCGGTACAGGG	2358	Db	3379	CCCCGATTCTGGTATAACCGGTATAACCGCTTGTGAGCTGATACCCGCTCGCGCA	5438
Qy	4075	GCTGGTTATCTGGTATAATCTGGACCGTGGCTGGGAGTCGGTACAGGG	4134	Qy	5155	GGCGAACGCGAGCCAGCAGTCAGTGGCTGGGAGTCGGTACAGGGCTCGGGT	5214
Db	2359	GCTGGTTATCTGGTATAATCTGGACCGTGGCTGGGAGTCGGTACAGGG	2418	Db	3439	GGCGAACGCGAGCCAGCAGTCAGTGGCTGGGAGTCGGTACAGGGCTCGGGT	5398
Qy	4135	CACTGCGCCAGGATAACAGATGCTGAGTGGTGGCTCACTGTTAACATT	4254	Qy	5215	ATTTCTCCCTGATGCTGAGTGGCTTACACTCCGCTATGGCTACGTCGGT	5274
Db	2419	CACTGCGCCAGGATAACAGATGCTGAGTGGCTTACCTGTTAACATT	2538	Db	3499	ATTTCTCCCTGATGCTGAGTGGCTTACACTCCGCTATGGCTACGTCGGT	5358
Qy	4195	CAACTATGATGAAACGAAATAGACAGATGCTGAGTGGTGGCTCACTGTTAACATT	4254	Qy	5275	TCTGCTGATGCTGCGCATAGTTAACGGCTTACACTCCGCTATGGCTACGTCGGT	5334
Db	2479	CAACTATGATGAAACGAAATAGACAGATGCTGAGTGGCTTACCTGTTAACATT	2538	Db	3559	TCTGCTGATGCTGCGCATAGTTAACGGCTTACACTCCGCTATGGCTACGTCGGT	3618
Qy	5335	CATGGCTGCGCCCGACACCCGGCAACCCGGCAACACCCGGCTGACGGCCTGAGA	5394	Qy			

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:	August 5, 2005, 18:42:20 ; Search time 3776 Seconds (without alignments)			
Title:	US-10-649-433C-1			
Perfect score:	6729			
Sequence:	1 gtttgcacatcttatcata... .agttagcgggattgtatctg 6729			
Scoring table:	Oligo_NUC			
Gapop:	60.0 , Gapext 60.0			
Searched:	7297361 seqs, 3241162794 residues			
Word size :	0			
Total number of hits satisfying chosen parameters:	14594722			
Minimum DB seq length:	0			
Maximum DB seq length:	200000000			
PPost-processing: Listing first 45 summaries				
Database :	Published Applications NA: 1: /cgn2_6/_ptodata/2/_pubpna/_us07_pubcomb.seq.* 2: /cgn2_6/_prodata/2/_pubpna/_us07_PCT_NEW_PUB.seq.* 3: /cgn2_6/_prodata/2/_pubpna/_us07__NEW_PUB.seq.* 4: /cgn2_6/_prodata/2/_pubpna/_us06__pubcomb.seq.* 5: /cgn2_6/_prodata/2/_pubpna/_us06__NEW_PUB.seq.* 6: /cgn2_6/_prodata/2/_pubpna/_us07_NEW_PUB.seq.* 7: /cgn2_6/_prodata/2/_pubpna/_us07_PCTUS_pubcomb.seq.* 8: /cgn2_6/_prodata/2/_pubpna/_us08__NEW_PUB.seq.* 9: /cgn2_6/_prodata/2/_pubpna/_us09A_pubcomb.seq.* 10: /cgn2_6/_prodata/2/_pubpna/_us09B_pubcomb.seq.* 11: /cgn2_6/_ptodata/2/_pubpna/_us09C_pubcomb.seq.* 12: /cgn2_6/_ptodata/2/_pubpna/_us09D_pubcomb.seq.* 13: /cgn2_6/_ptodata/2/_pubpna/_us10A_pubcomb.seq.* 14: /cgn2_6/_ptodata/2/_pubpna/_us10B_pubcomb.seq.* 15: /cgn2_6/_ptodata/2/_pubpna/_us10C_pubcomb.seq.* 16: /cgn2_6/_ptodata/2/_pubpna/_us10D_pubcomb.seq.* 17: /cgn2_6/_ptodata/2/_pubpna/_us10E_pubcomb.seq.* 18: /cgn2_6/_ptodata/2/_pubpna/_us10F_pubcomb.seq.* 19: /cgn2_6/_ptodata/2/_pubpna/_us10G_pubcomb.seq.* 20: /cgn2_6/_ptodata/2/_pubpna/_us10H_pubcomb.seq.* 21: /cgn2_6/_ptodata/2/_pubpna/_us10I_pubcomb.seq.* 22: /cgn2_6/_ptodata/2/_pubpna/_us10J_NEW_PUB.seq.* 23: /cgn2_6/_ptodata/2/_pubpna/_us11A_pubcomb.seq.* 24: /cgn2_6/_ptodata/2/_pubpna/_us11B_NEW_PUB.seq.* 25: /cgn2_6/_ptodata/2/_pubpna/_us60_NEW_PUB.seq.* 26: /cgn2_6/_ptodata/2/_pubpna/_us60__pubcomb.seq.*			
SUMMARIES				
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				
Result No.	Score	Query Match Length	DB ID	Description
1	6729	100.0	6729	21 US-10-649-433C-1
2	3804	56.5	5013	14 US-10-198-03A-5
3	3804	56.5	5013	15 US-10-198-03B-5
c 4	3753	55.8	5926	13 US-10-024-809-3
c 5	3753	55.8	5926	13 US-10-024-809-4
c 6	3702	55.0	4176	16 US-10-124-1424A-3
c 7	2741	40.7	5772	20 US-10-131-998A-17

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	ID			
1	6729	100.0	6729	21	US-10-649-433C-1		Sequence 1, Appli
2	3804	56.5	5013	14	US-10-198-034-5		Sequence 5, Appli
3	3804	56.5	5013	15	US-10-198-228-5		Sequence 5, Appli
c 4	3753	55.8	5926	13	US-10-024-009-3		Sequence 3, Appli
c 5	3753	55.8	5926	13	US-10-024-809-4		Sequence 4, Appli
6	3702	55.0	4176	16	US-10-244-142A-3		Sequence 3, Appli
7	2741	40.7	5772	20	US-10-131-098A-17		Sequence 37, Appli

```

Query Match      100 %;  Score 6729;  DB 21;  Length 6729;
Best Local Similarity 100 %;  Pred. No. 0;
Matches 6729;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

1 GTTGAGCTTATCATGACTGGCACCAATGCTTCTGGGTAGGCATC 60
   ||||| | | | | | | | | | | | | | | | | | | | | | | |
1 GTTGAGCTTATCATGACTGGCACCAATGCTTCTGGGTAGGCATC 60
   ||||| | | | | | | | | | | | | | | | | | | | | | |
61 GAAAGCTGTGGTGTGGGTGAATCATCTGACCTGGCCATTGCTTCAGGC 120
   ||||| | | | | | | | | | | | | | | | | | | | | | |
61 GAAAGCTGTGGTGTGGGTGAATCATCTGACCTGGCCATTGCTTCAGGC 120
   ||||| | | | | | | | | | | | | | | | | | | | | |

```

Qy	121	GCACTCCGGTTCTGGATAATGTTTTCGGCGACATCATAAAGGGTTCTGGCAAATATTC	180	Db	1201	GGCTATTGGAAAGGTGGAAAGGCCGCCAGATTTCAAGCTGCGATGATGATGTTCTCA	1260
Db	121	GCATCCTCGTTCTGGATAATGTTTTCGGCGACATCATAAAGGGTTCTGGCAAATATTC	180	Qy	1261	GAATGACCACTTCCAAAGCAGACGATCAGATGACCTTATTCGGCCAAGGGGTGAAAGC	1320
Qy	181	TGAAATGGCTGTTGACATTAATCATCGGTCGATTAATGTTGCGAATGTTGCGGA	240	Db	1261	GAATGACCACTTCCAAAGCAGACGATCAGATGACCTTATTCGGCCAAGGGGTGAAAGC	1320
Db	181	TGAAATGGCTGTTGACATTAATCATCGGTCGATTAATGTTGCGAATGTTGCGGA	240	Qy	1321	ACTGGCCTTACACCGCTTACCCCGACAGTCGGGTTACCGTGAATGAGAAAGCCGTTG	1380
Qy	241	TAAACAATTTCACAGGAAACAGGGGCTGTGAGAAAAGCSAAGGGCACTGCTTTAA	300	Db	1321	ACTGGCCTTACACCGCTTACCCCGACAGTCGGGTTACCGTGAATGAGAAAGCCGTTG	1380
Db	241	TAAACAATTTCACAGGAAACAGGGGCTGTGAGAAAAGCSAAGGGCACTGCTTTAA	300	Qy	1381	GCAAAAGTGCGGGGTTCTCAACAGAACGCTGCTGAGGCTGATAGCTAAGCTGTA	1440
Qy	301	CAATTATCAGACAATCTGTTGGCGACACTCGACCGGAAATTATCGATACTTATTATA	360	Db	1381	GCAGAACATGCGGGGTTCTCAACAGAACGCTGCTGAGGCTGATAGCTAAGCTGTA	1440
Db	301	CAATTATCAGACAATCTGTTGGCGACACTCGACCGGAAATTATCGATACTTATTATA	360	Qy	1441	CGACAAAGCCCTACTAGTTGECACTGACTCAAAAGTAGTCCGGCATATTCAAGGGATT	1500
Qy	361	AAAATTAAAGAGTATATTAATGTTATGTTATGTTAAACCATGGTAG	420	Db	1441	CGACAAAGCCCTACTAGTTGECACTGACTCAAAAGTAGTCCGGCATATTCAAGGGATT	1500
Db	361	AAAATTAAAGAGTATATTAATGTTATGTTAAACCATGGTAG	420	Qy	1501	GATTGCTAAACACTGGGGGGAATCACGGGTTGGATCTGAGCTCAAAAGACGCTCAGATTCA	1560
Qy	421	CAAGGGCAAGGAGCTGTTACCGGGTTGTGCCATCTGGTGCACGGCCACGT	480	Db	1501	GATTGCTAAACACTGGGGGGAATCACGGGTTGGATCTGAGCTCAAAAGACGCTCAGATTCA	1560
Db	421	CAAGGGCAAGGAGCTGTTACCGGGTTGTGCCATCTGGTGCACGGCCACGT	480	Qy	1561	GTTTCGTTACTGTGTGAAGGTGACCGCTTACAGTTAGTAACTCCAAAT	1620
Qy	481	AAACGGCCACAAAGTTCAGGTTTCGGGCAACTACGGCGAACGT	540	Db	1561	GTTTCGTTACTGTGTGAAGGTGACCGCTTACAGTTAGTAACTCCAAAT	1620
Db	481	AAACGGCCACAAAGTTCAGGTTTCGGGCAACTACGGCGAACGT	540	Qy	1621	GATTAAAGAATTGAAAGGATTAAGGGCATCAAACCTGAACTAGTTAGTAACTCCAA	1680
Qy	541	GACCCTGAAGTTCATCAGACCAACCCGAAAGTGGCGCTGCCAACCTCTGAC	600	Db	1621	GATTAAAGAATTGAAAGTAACTGATGAACTGATGAACTGATGAACTCCAA	1680
Db	541	GACCCTGAAGTTCATCAGACCAACCCGAAAGTGGCGCTGCCAACCTCTGAC	600	Qy	1681	GTGGGACACCCCTCAGGGAAAGATAAGATGAGCTGCGCTGCGATGCGCTGCGAAGCGCA	1740
Qy	601	CACCTTCGCTTACGGCGTCACTGGCGTCACTGGCGCTTACCCGACACATGAGGAGCA	660	Db	1681	GTGGGACACCCCTCAGGGAAAGATAAGATGAGCTGCGCTGCGATGCGCTGCGAAGCGCA	1740
Db	601	CACCTTCGCTTACGGCGTCACTGGCGTCACTGGCGCTTACCCGACACATGAGGAGCA	660	Qy	1741	AAAAATGGAATGGTTATGCCAACACGATGCGCTGCGATGCGCTGCGAAGCGCA	1800
Qy	661	CTTCCTGAAGTCGCCATGCGGCTACCCGACCATCTCTTCAGGA	720	Db	1741	CAAAATGGAATGGTTATGCCAACACGATGCGCTGCGATGCGCTGCGAAGCGCA	1800
Db	661	CTTCCTGAAGTCGCCATGCGGCTACCCGACCATCTCTTCAGGA	720	Qy	1801	GAAGGGACACACAAAGTCCACGATTCGGGTTTGGCGTCAAGAGCGCT	1860
Qy	721	CGACGGCAACTACAGACGGCCGAGCTGGGCAACCCCTGGTGAACCG	780	Db	1801	GAAGGGACACACAAAGTCCACGATTCGGGTTTGGCGTCAAGAGCGCT	1860
Db	721	CGACGGCAACTACAGACGGCCGAGCTGGGCAACCCCTGGTGAACCG	780	Qy	1861	GGCGCTGGTGAATGGTGAACCTGGGCACTGGGCAACCCCTGGGCAACTGGGCAACCC	1920
Qy	781	CATCGAGGTGAAGGGCATGACTTCAGGAGACGGAAACATTCTGGGCAAGCTGGA	840	Db	1861	GGCGCTGGTGAATGGTGAACCTGGGCACTGGGCAACCCCTGGGCAACTGGGCAACCC	1920
Db	781	CATCGAGGTGAAGGGCATGACTTCAGGAGACGGAAACATCTGGGCAAGCTGGA	840	Qy	1921	GAAGGGACACTTGTGACTGGGCAAAACCTGGGCAATGGTGGCTGATGGCAC	1980
Qy	841	GTACAACATACAGCCCAACTCTATATCATGGGCCACAGGAGCGCATCAA	900	Db	1921	GAAGGGACACTTGTGACTGGGCAAAACCTGGGCAATGGTGGCTGATGGCAC	1980
Db	841	GTACAACATACAGCCCAACTCTATATCATGGGCCACAGGAGCGCATCAA	900	Qy	2041	CCTGGTGTGAATGGTGAACCTGGGCAAAACTGGGCAATGGTGGCTGATGGCAC	2100
Qy	901	GTTGAACATCAAGATCCGCCACACATGAGGACGGCAACCTGGGCAACACTA	960	Db	2041	CCTGGTGTGAATGGTGAACCTGGGCAAAACTGGGCAATGGTGGCTGATGGCAC	2100
Db	901	GTTGAACATCAAGATCCGCCACACATGAGGACGGCAACCTGGGCAACACTA	960	Qy	2101	CCCAATTCTGGTGAATGGTGAACCTGGGCAAAACTGGGCAATGGTGGCTGATGGCAC	2160
Qy	961	CCAGCAGAACCCCCATGGCAAGGGCCCTGGCTGCGCAACCTACCTGG	1020	Db	2101	CCCAATTCTGGTGAATGGTGAACCTGGGCAAAACTGGGCAATGGTGGCTGATGGCAC	2160
Db	961	CCAGCAGAACCCCCATGGCAAGGGCCCTGGCTGCGCAACCTACCTGG	1020	Qy	2161	GGGTGAGGTGATGCAACATGGGAAACTTCTGCAATGGGAAACTTCTGCAATGGT	2220
Qy	1021	CTACCACTGGCCCTGAGCAAAGCCCCAACGAGGAAGGGGATCACATGGCTGGA	1080	Db	2161	GGGTGAGGTGATGCAACATGGGAAACTTCTGCAATGGGAAACTTCTGCAATGGT	2220
Db	1021	CTACCACTGGCCCTGAGCAAAGCCCCAACGAGGAAGGGGATCACATGGCTGGA	1080	Qy	2221	ACTACCTGTTCCATGGCCAAACCTGGTCACTACTTCTGTTATGGTGTCAATGGTTTC	2280
Qy	1141	TACTCGCTTGGTGTAAACATCTAACGAGATAACTTATGTTGAGTGCCTA	1200	Db	2221	ACTACCTGTTCCATGGCCAAACCTGGTCACTACTTCTGTTATGGTGTCAATGGTTTC	2280
Db	1141	TACTCGCTTGGTGTAAACATCTAACGAGATAACTTATGTTGAGTGCCTA	1200	Qy	2281	CGGTATCCGGATCATATGAAACGGGATGACTTTTCAGGTGGCCGAAAGCTTA	2340

Result No.	Score	Query	Match	Length	DB ID	Description
1	309	14.2	332	1	JGBCG	D-Galactose-bindin galactose-binding lactose operon rep lac repressor - sy
2	309	14.2	332	2	B91009	beta-lactamase (BC)
3	309	14.2	332	2	D85853	beta-lactamase (BC)
4	284	13.0	360	1	RPECL	beta-lactamase (BC)
5	284	13.0	360	4	S58116	beta-lactamase (BC)
6	204	9.4	286	2	T51301	beta-lactamase (BC)
7	204	9.4	286	2	S47061	beta-lactamase (BC)
8	204	9.4	286	4	S41975	beta-lactamase (BC)
9	185	8.5	286	1	PNECP	beta-lactamase (BC)
10	185	8.5	286	4	I40305	beta-lactamase (BC)
11	183	8.4	360	2	B85529	beta-lactamase (BC)
12	183	8.4	360	2	P96578	beta-lactamase (BC)
13	104	4.8	105	2	JC2566	beta-lactamase (BC)
14	104	4.8	286	2	S60312	extended spectrum beta-lactamase (EC)

Result No.	Score	Query Match	Length	DB ID	Description
c 1	864	12.8	1070	1 AJ281552	AJ281552 4A3A-P671
c 2	788	11.7	841	1 AL042026	AL042026 DKF2P44E
c 3	742	11.4	865	7 CK125894	CK125894 BES182411
c 4	742	11.0	793	5 BQ751655	BQ751655 EST732218
5	742	11.0	806	5 BQ751220	BQ751220 EST7631783
c 6	742	11.0	827	7 CN8233902	CN8233902 Oa_sp1bn
c 7	741	11.0	758	9 CL422788	CL422788 AE0544_Sa
c 8	736	10.9	856	7 CN823189	CN823189 Oa_sp1bn
c 9	732	10.9	928	7 C0487414	C0487414 G00227_B7
c 10	723	10.7	1004	1 AJ281480	AJ281480 4A3A-P4G8
c 11	722	10.7	789	6 CD280920	CD280920 G4224_42
c 12	717	10.7	769	7 CK118014	CK118014 218n06_P1
c 13	706	10.5	759	6 CD279661	CD279661 G4318_35
c 14	699	10.4	780	5 BQ845693	BQ845693 101012980
c 15	699	10.4	846	7 CV468077	CV468077 est.1 van
c 16	690	10.3	741	6 CD279174	CD279174 G4221_93
c 17	679	10.1	730	6 CD281097	CD281097 G4224_38
c 18	678	10.1	729	6 CD279546	CD279546 G4222_81
c 19	678	10.1	800	1 AJ281449	AJ281449 4A3A-P4D5
c 20	677	10.1	728	6 CD279322	CD279322 G4223_13
c 21	675	10.0	726	6 CD81811	CD81811 G4471_30
c 22	674	10.0	954	1 AL044364	AL044364 DKFPz434C
c 23	673	10.0	710	7 CK94569	CK94569 hggada540
c 24	671	10.0	733	8 BZ049372	BZ049372 jnr50b10.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

ON nucleic - protein search, using frame_plus_n2p model

Run on: August 2, 2005, 19:25:21 ; Search time 1028 Seconds

(without alignments)
6703.854 Million cell updates/sec

Title: US-10-649-433C-1

Perfect score: 2180

Sequence: 1 gttgcacgcttatcatcgaaaaaaaaatgttagcgccattatgtatctg 6729

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delet 7.0

Searched: 1612378 seqs., 512079187 residues

Word size: 1

Total number of hits satisfying chosen parameters:

3224408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p model -DEV=x1D
-Q=/cgn/1/USPRO_spool_P/US10649433/runat_02082005_165025_1820/app_query.fasta_1.6919
-DB=UniProt -QFMT=fasian -SUFFIX=_01go.rup -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TPANS=human10 cdi -LIST=45
-DOCALLIGN=1 -THR SCORE=qrunat_02082005_165025_1820 -NCPU=6 -ICPU=3
-NORM=ext -HBAPSIZE=500 -MINLEN=2000000000
-USERID=10619433 @CGN 1 1 1271 @runat_02082005_165025_1820 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPPOP=60 -XGAPEXT=60 -FGAPPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELPOP=6 -DELEXT=7
```

Database : UniProt 03.1

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	3.09	14.2	332	1	DGAL_ECOLI	P02927 escherichia	P02927 escherichia
2	3.09	14.2	332	2	Q8X641	escherichia	Q8X641 escherichia
3	2.84	13.0	360	1	LACI_ECOLI	P03023 escherichia	P03023 escherichia
4	2.22	10.2	286	2	Q9IUY8	plasmid pww	Q9IUY8 plasmid pww
5	2.04	9.4	286	2	Q38058	bacteriophaga	Q38058 bacteriophaga
6	2.04	9.4	286	2	Q00626	staphylococci	Q00626 staphylococci
7	2.04	9.4	286	2	Q79CL6	methylohaci	Q79CL6 methylohaci
8	2.04	9.4	286	2	Q79DR3	escherichia	Q79DR3 escherichia
9	2.03	9.3	285	2	Q6KB67	hordeum vulgare	Q6kb67 hordeum vulgare
10	1.95	8.5	286	1	BLAT_ECOLI	P62593 escherichia	P62593 escherichia
11	1.85	8.5	286	1	BLAT_SALTII	P62594 salmonella	P62594 salmonella
12	1.85	8.5	286	1	Q6A253	haemophilus	Q6a253 haemophilus
13	1.85	8.5	286	2	Q6LB9N	pseudomonas	Q6lb9n pseudomonas
14	1.85	8.5	286	2	Q6LCV6	neisseria	Q6lcv6 neisseria
15	1.85	8.5	286	2	Q6TMH1	streptococcus	Q6tmh1 streptococcus
16	1.85	8.5	286	2	Q6W9J1	enterobacteria	Q6w9j1 enterobacteria

ALIGMENTS

RESULT 1	ID	DGAL_ECOLI	STANDARD	PRT;	332 AA.
AC	P02927; P17775;				
DT	21-JUL-1987 (Rel. 01, Created)				
DT	13-AUG-1987 (Rel. 05, Last sequence update)				
DT	25-JAN-2005 (Rel. 46, Last annotation update)				
DE	D-Galactose-binding periplasmic protein precursor (GBP) (D-galactose-binding protein) (GGP)				
GN	Name=mglB; OrderedLocusNames=d2150, c2684, SF2235, S2364;				
OS	Escherichia coli O6, and				
OS	Shigella flexneri.				
OC	Bacteri; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
NCBI_TAXID	562; 217992, 623;				
RN	"Nucleotide sequence and analysis of the mgl operon of Escherichia coli K12."				
RC	SEQUENCE FROM N.A.				
RC	SPECIES=E.coli; STRAIN=K12;				
RX	MEDLINE=92049246; PubMed=1119366;				
RA	HOGG R.W., Voecker C., von Carlowitz I.;				
RA	"Nucleotide sequence and analysis of the mgl operon of Escherichia coli K12."				
RT	SEQUENCE FROM N.A.				
RT	SPECIES=E.coli; STRAIN=K12;				
RT	SEQUENCE FROM N.A.				
RT	RICHARDICH P., LAKEY N., GRYAN G., JAEHN L., MINTZ L., ROBISON K.,				
RA	CHURCH G.M.;				
RA	Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.				
RN	SEQUENCE FROM N.A.				
RP	SEQUENCE FROM N.A.				

17	185	8.5	286	2	Q6W9J2	zymomonas m
18	185 <td>8.5 <td>286</td> <th>2</th> <th>Q6WZD4</th> <th>acetinobact</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q6WZD4</th> <th>acetinobact</th>	286	2	Q6WZD4	acetinobact
19	185 <td>8.5 <td>286</td> <th>2</th> <th>Q733X5</th> <th>citrrobacter</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q733X5</th> <th>citrrobacter</th>	286	2	Q733X5	citrrobacter
20	185 <td>8.5 <td>286</td> <th>2</th> <th>Q7B899</th> <th>klebsiella</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q7B899</th> <th>klebsiella</th>	286	2	Q7B899	klebsiella
21	185 <td>8.5 <td>286</td> <th>2</th> <th>Q7BP57</th> <th>shigella fl</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q7BP57</th> <th>shigella fl</th>	286	2	Q7BP57	shigella fl
22	185 <td>8.5 <td>286</td> <th>2</th> <th>Q7BR75</th> <th>neisseria m</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q7BR75</th> <th>neisseria m</th>	286	2	Q7BR75	neisseria m
23	185 <td>8.5 <td>286</td> <th>2</th> <th>Q7DYY3</th> <th>salmonella m</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q7DYY3</th> <th>salmonella m</th>	286	2	Q7DYY3	salmonella m
24	185 <td>8.5 <td>286</td> <th>2</th> <th>Q7DH33</th> <th>serratia ma</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q7DH33</th> <th>serratia ma</th>	286	2	Q7DH33	serratia ma
25	185 <td>8.5 <td>286</td> <th>2</th> <th>Q799Y1</th> <th>plasmid ppv</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q799Y1</th> <th>plasmid ppv</th>	286	2	Q799Y1	plasmid ppv
26	185 <td>8.5 <td>286</td> <th>2</th> <th>Q7BV8</th> <th>plasmid pet</th> </td>	8.5 <td>286</td> <th>2</th> <th>Q7BV8</th> <th>plasmid pet</th>	286	2	Q7BV8	plasmid pet
27	183 <td>8.4 <td>286</td> <th>2</th> <th>Q7AH57</th> <th>escherichia</th> </td>	8.4 <td>286</td> <th>2</th> <th>Q7AH57</th> <th>escherichia</th>	286	2	Q7AH57	escherichia
28	183 <td>8.4 <td>286</td> <th>2</th> <th>Q8x684</th> <th>escherichia</th> </td>	8.4 <td>286</td> <th>2</th> <th>Q8x684</th> <th>escherichia</th>	286	2	Q8x684	escherichia
29	181	8.3	225	2	Q38212	bacteriophag
30	181	8.3	281	2	Q6avm7	acinetobact
31	181	8.3	282	2	Q6W7J4	escherichia
32	181	8.3	286	2	Q8XSD3	klebsiella
33	181	8.3	286	2	Q93A77	escherichia
34	181	8.3	286	2	Q93G13	klebsiella
35	181	8.3	286	2	Q6SQJ9	streptococc
36	181	8.3	286	2	Q9R771	klebsiella
37	181	8.3	286	2	Q6W7J4	escherichia
38	181	8.3	286	2	Q8GA85	escherichia
39	181	8.3	286	2	Q9EYX1	escherichia
40	181	8.3	286	2	Q9E582	klebsiella
41	181	8.3	286	2	Q9R771	klebsiella
42	181	8.3	286	2	Q9RMS4	escherichia
43	181	8.3	286	2	Q9RN48	escherichia
44	181	8.3	286	2	Q93679	escherichia
45	180	8.3	286	2	Q93328	escherichia

RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=9426617; PubMed=278503; DOI=10.1126/science.277.5331.1453;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Billey M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.; "The complete genome sequence of *Escherichia coli* K-12.";
Science 277:1453-1474(1997).
[5]

RL SEQUENCE FROM N.A.
RN SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=23388234; PubMed=12471157; DOI=10.1073/pnas.252520799;
Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J.T., Stroud D.,
Mayhew G.F., Rose J., Zhou S., Schwartz D.C., Perna N.T.,
Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
"Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic *Escherichia coli*.",
Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
[6]

RN SEQUENCE OF 24-332.
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=81168234; PubMed=7012152;
RA Mahoney W.C., Hogg R.W., Hermodson M.A.;
RA "The amino acid sequence of the D-galactose-binding protein from
Escherichia coli B/r.";
RT J. Biol. Chem. 256:4350-4356 (1981).
[7]

RN REVISIONS TO 179-180.
RP SEQUENCE=E.coli;
RC SPECIES=E.coli;
RX MEDLINE=8291030; PubMed=6885805;
RA Scripture J.B., Hogg R.W.;
RT "the nucleotide sequences defining the signal peptides of the
galactose-binding protein and the arabinose-binding protein.";
RT J. Biol. Chem. 258:10853-10855 (1983).
[9]

RN SEQUENCE OF 24-47.
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=K12 / ENRG2;
RX MEDLINE=9443975; PubMed=298646;
RA Link A.J., Robison K., Church G.M.;
RA "Comparing the predicted and observed properties of proteins encoded
in the genome of *Escherichia coli* K-12.";
RT Electrophoresis 18:1259-1313 (1997).
[10]

RN SEQUENCE FROM N.A.
RP SEQUENCE=S.flexneri; STRAIN=301 / Serotype 2a;
RC SPECIES=S.flexneri; STRAIN=2457 / ATCC 700930 / Serotype 2a;
RX MEDLINE=2272405; PubMed=12384550; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang F., Yang J., Yang X., Zhang Y., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao J., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RA "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
through comparison with genomes of *Escherichia coli* K12 and O157.";
RT Nucleic Acids Res. 30:4432-4441 (2002).
[11]

RN SEQUENCE FROM N.A.
RP SEQUENCE=S.flexneri; STRAIN=2457 / ATCC 700930 / Serotype 2a;
RC SPECIES=S.flexneri; STRAIN=2590274; PubMed=12704152;
RX MEDLINE=2590274; PubMed=12704152;
DOI=10.1126/IAI.71.5.2775-2786 (2003);
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
Mau B., Perna N.T., Payne S.M., Runyan-Janecky L.J., Zhou S.,
Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of *Shigella*
flexneri serotype 2a strain 2457.";
Infec. Immun. 71:2775-2786 (2003).
[12]

GenCore version 5.1.6	GenCore search, using frame_plus n2p mode	Adh14282 Vector PC
Copyright (c) 1993 - 2005 Compugen Ltd.		Adh45227 Modified
I nucleic - protein search, using frame_plus n2p mode		Adh14284 Vector pc
on on:	August 2, 2005, 18:10:32 : Search time 802 Seconds (without alignments) 6490.051 Million cell updates/sec	Aar07640 Deduced p
title:	US-10-649-433C-1	Aar04031 Full leng
perfect score:	11920	Aar52701 Plasmid
quence:	1 gtttagacagcttatcatcgaaaaaaaaatgtatcgatctg 6729	Adg15745 Novel vac
oring table:	BLOSUM62	Aar52700 Plasmid P
Xgapop 10.0 , Xgapext 0.5	C 11 3004 24.9	Adc22808 Human G P
Ygapop 10.0 , Ygapext 0.5	C 12 2946.5 24.7	Adh14282 Vector pc
Fgapop 6.0 , Fgapext 7.0	C 13 2902 24.3	Adc22811 Human G P
Delop 6.0 , Delext 7.0	C 14 2902 24.3	Adh14284 Vector pc
arched:	2105692 seqs, 386766381 residues	Aar07640 Deduced p
total number of hits satisfying chosen parameters:	4211384	Aar04031 Full leng
minimum DB seq length: 0	C 15 2801.5 24.1	Aar07640 Deduced p
maximum DB seq length: 2000000000	C 16 2801.5 24.1	Aar07641 Deduced s
st-processing: Minimum Match 100%	C 17 2895.5 24.1	Aar07641 Deduced s
Maximum Match 100%	C 18 2890.5 24.0	Aar07641 Deduced s
Listing first 45 summaries	C 19 2871 24.1	Aar07641 Deduced s
command line parameters:	C 20 2842 23.6	Aar07641 Deduced s
ODEL=frame+n2p.model -DEV=xlp	C 21 2842 23.6	Aar07641 Deduced s
UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi -LIST=45	C 22 2824 23.7	ADH14285 Vector PC
OCALIGN=200 -THR MIN=0 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL	C 23 2824 23.7	Adh14286 Vector PC
UTEMT=pco -NORM=ext	C 24 2823 23.7	Aar52702 Plasmid P
ISER=US10649433 @CGRN_1_977 @runat 02082005.164944.1513 -NCPU=6 -ICPU=3	C 25 2819 23.4	Adh14283 Vector PC
O MMAP -LARGEQUERY -NNEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG	C 26 2819 23.4	Aar07641 Deduced s
TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6	C 27 2802 23.2	AAR07641 Deduced s
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7	C 28 2773.5 23.3	ADC22813 Human G P
	C 29 2773.5 23.3	Adh14286 Vector PC
	C 30 2709 22.7	Aar52702 Plasmid P
	C 31 2707.5 22.5	Adh14283 Vector PC
	C 32 2675.5 22.4	Aar07641 Deduced s
	C 33 2626 22.0	Adh11230 Vertebrat
	C 34 2618.5 22.0	Aar88634 Plasmid P
	C 35 2617 22.0	Abr13387 FlIPrbs f
	C 36 2592 21.7	ADH14287 Vertebrat
	C 37 2586 21.7	Adh11228 Vertebrat
	C 38 2486.5 20.9	Aab93360 Human H-R
	C 39 2486.5 20.9	Aab9357 Human H-R
	C 40 2486.5 20.9	Abb06725 Human H-R
	C 41 2486.5 20.9	Abb06728 Human H-R
	C 42 2486.5 20.9	Abb06738 Human H-R
	C 43 2486.5 20.9	Abb06735 Human H-R
	C 44 2486.5 20.9	Abb06730 Human H-R
	C 45 2486.5 20.9	Abb06740 Human H-R
		RESULTS
		1 ABB82793
		ID ABB82793 standard; protein; 1967 AA.
		AC ABB82793;
		XX
		DT 18-MAR-2003 (first entry)
		XX
		DB Amino acid sequence of plasmid pIrcCFRET3.
		XX
		XX Fluorescent; fluorescence resonance energy transfer; RET; tumour;
		KW estrogen; tamoxifen; therapy; pIrcCFRET3.
		XX Synthetic.
		OS
		XX
		XX Key Location/Qualifiers
		XX Misc-difference 1 . 1987
		XX /note= "Xaa are residues encoded by internal stop codons"
		XX
		XX WO200290987-A2.
		XX
		XX PN
		XX PD 14-NOV-2002.
		XX PP 10-MAY-2002; 2002WO-GB002183.
		XX PR 10-MAY-2001; 2001GB-00011459.
		XX PA (ISIS-) ISIS INNOVATION LTD.
		XX PI Fricker MD, Vaux DJT;
		XX DR WPI; 2003-129191/12.
		XX DR N-PSDB; ABZ23939.
		XX XX
		SUMMARIES
		§
atabase :	A_Geneseq_16Dec04: 1: geneseqP1980s: 2: geneseqP1990s: 3: geneseqP2000s: 4: geneseqP2001s: 5: geneseqP2002s: 6: geneseqP2003as: 7: geneseqP2003bs: 8: geneseqP2004s: 9: geneseqP2005s: 10: geneseqP2006s:	Result No. Score Match Length DB ID Description
1	6839.5 72.5 1967 6 ABB82793 Abb88793 Amino aci	
2	40.04 33.6 799 6 ABR43385 Abr13385 FLIPrbs f	
3	3273 27.1 2606 7 ADB52217 Ade2217 PG1EN-EH3	
4	3273 27.1 2606 7 ADB52214 Ade2214 PG1EN-EH3	
5	3086.5 25.9 2761 2 ADH11249 Ade11249 Vertebrat	
6	3083.5 25.6 3614 2 ADH11258 Ade11258 Vertebrat	
7	3062.5 25.7 766 2 ADH11236 Ade11236 Vertebrat	
8	3058 25.7 1221 2 AAR52639 Aar52639 Sequence	
9	3054 25.7 2732 2 ADH11251 Ade11251 Vertebrat	
10	3004 24.9 1310 7 ADC22809 Aar22809 Human G P	

PT Probe, useful e.g. for medical diagnosis, and detection of pollutants in water systems and contaminants in foodstuffs, has target binding site PT moiety and mimic moiety attached to two fluorescent polypeptides, and PT linker.

XX Disclosure; Fig 5A-H; 50pp; English.

PS The invention relates to a probe (I) comprising a target binding site CC capable of binding to a first fluorescent polypeptide, a mimic moiety CC capable of binding to the target binding site moiety and attached to a second fluorescent polypeptide, and a linker connecting the two CC fluorescent polypeptides and which allows the distance between the CC polypeptides to vary. (I), or a cell harbouring (I), or a light CC polynucleotide which encodes (I), or a sensor (IV) comprising (I), a light source which is capable of exciting the probe, and a detecting which is CC capable of measuring the amount of fluorescence resonance energy transfer (FRET) from the probe are useful for detecting the presence or absence of CC a target substance in a test sample, and for identifying an inhibitor of CC binding between two substances, where the two substances would bind to each other in the absence of an inhibitor (see AB223939 for a description CC of the varied uses of the probe of the invention). The present sequence XX represents the amino acid sequence of plasmid pTRCFRET3

SQ Sequence 1967 AA;

Alignment Scores:

Pred. No. :	0	Length:	1967	Db	184	AspHisTyrGlnGlnAsnThrProIleGlyAspGlyProValLeuProAspAsnHis 203
Score:	8639.50	Matches:	1701	Qy	1013	TACCTAGTCACCGTCCGCCCTGGCAANGACCCAACCAGAACGGGATCACATGGTC 1072
Percent Similarity:	81.30%	Conservative:	21	Db	204	TyrLeuSerIleGlnSerAlaLeuSerIleAspProAsnGluYargAspHisLeu 223
Best Local Similarity:	80.31%	Mismatches:	76	Qy	1073	CTGCTGAGTTGTCGACCGGCC-----GGCATCTACTCTC 1108
Query Match:	72.48%	Indels:	320	Db	224	LeuLeuGluPheValIleAlaLeuGlnSerIleSerIleAsp-----LeuGlySerIleMet 243
DB:	6	Gaps:	8	Qy	1109	GGCATGGACGAG-----CTGTAAGACTAGTGTACTGCATT 1150
US-10-649-433C-1 (1-729) x ABB82793 (1-1967)		Db	244	GlyGlySerIleHisHisGlyMetAlaSerMetThrGlyGlnGlnMet 263		
Qy	413	ATGGTGAGCAAGGGCGAGGAGCTGTTCAACCGGGTCTGGCCATCTGGTCGAGCTGAC 472	Qy	1151	GGTGTAACAAVTAATAGTAGACATAACTTATGTCCTAGTCGCAAGGCTATTGAG 1210	
Db	4	MetValSerIleGlyGluGluLeuPheThrGlyValValProIleLeuAla 23	Db	264	GlyArgAspLeuTyroSpaAspAspSp----- 272	
Qy	473	GCGGACCTAAAGGCCAACAGTTACGGGTGTCGGCAAGGGAGGAGGCTACCTAC 532	Qy	1211	CAAGATGCGAAAAGCCCAGCAAGATTCAGCTGCTGATGAATGATACTGAGGACTGACCAG 1270	
Db	24	GlyAspValAargIlyHiSlysPheSerIleGlyGluGlyGluGlyGlu 43	Db	272	----- 272	
Qy	533	GCGAACGCTGACCCCTGARGTTCATCTGCACCAACGCCAACAGCTGCCCTGGCCACC 592	Qy	1271	TCCAAGCAGAACGATCAAGATGACCTAGTCGTTGCCCCAGGACTGCCCAT 1330	
Db	4	GlyLysIleIleThrIleLeuIlePheThrGlyIleLeuProValProIleProThr 63	Db	272	----- 272	
Qy	593	CTCGTGCACACCTTCGGCTACGGCCATGCTGCTGGCCCTAACCGGACACATGAG 652	Qy	1331	AACCTGGTTGACCGCACGACTCAAAAGAGTCGGCATTTGAGGAGCTGGCAAAAC 1390	
Db	64	IleLeuIleThrIleLeuIleTrpGlyValGlnGlySerIleAspAlaThrIle 43	Db	272	----- 272	
Qy	653	CAGCACCACTCTTCAAGTCCCACATGCCAGGCTAAGTCGAGAGGCTAAGTC 712	Qy	1451	TACTAGTTGGCACIGACTCAAAAGAGTCGGCATTTGAGGAGCTGGCTGAAAC 1510	
Db	84	GlnHiBspPhePhelySerAlaIleProGluGlyIleLeuGlyGlnGly 103	Db	273	----- 273	
Qy	713	TTCAGGACGAGCCGAACTACAAGACCGCCGCAACTGAACTGGGGGACACCTG 772	Qy	1511	CACTGGGGGCCATTAGGGTTGGATCTGACATGAGCTGGCTACTG 1570	
Db	104	PhoLysAspBspGlyAsnTyroSthArgAlaGluIlePheGluYaspThrLeu 123	Db	274	----- 274	
Qy	773	GtgAACCGCATCGAGCTGAAAGGCATCGACTCAAGGGACGGCAACATCTGGGCAC 832	Qy	1571	CTGAAGTGTGACCGGCCATCGGATGCAAGGAGCTAACCTAGCTGATTAAAGAA 1630	
Db	124	ValAlaArgIleIleLeuIleGlyIleAspPhoLysAsnIleLeuGlyHis 143	Db	279	ProArgGlySerSerIle 284	
Qy	833	AAGCTGAGTGTACACTAACGCAACGCTTATCATGGCCACACAGGAGAAC 892	Db	1631	TTGAAAGTAAAGGCAAAACTGATGAACTTACACTTAGATACTGCAATGGGACACC 1690	
Db	144	LysLeuGluTyroAsnTyroIleSerHisAsnValTyzIleThrAlaAspLysGln 163	Qy	1691	GCTCTGGGAAAGATAAGATGGACCTGCTGGCTGCCCCGAAACGCCAACATCGAA 1750	
Qy	893	GGCATCAAGGTGAACTTAAAGTCGCAACACATCGGAAGGCAAGCTGCGCC 952	Db	284	----- 284	
Db	164	GlyIleLysAlaAsnPhoLysIleArgIleIleLeuIleGlySerIleGln 183	Qy	1751	GTGGTTATCGCCAAACAGATGCGATGCAATGGGGCGGTTGAAGGCTCAAAGCACAC 1810	
Qy	953	GACCACTTACCGAACCCCCATCGGCCAGCTGCTGCTGCTGCTGCTGCTGCTG 1012	Db	284	----- 284	
Db	296	----- 296	Qy	1811	AAACAGTCCAGCATTCGGTTTGGCTGATGGCTGCTGGCGAACGGCAAC 1870	
Qy	1811	AAACAGTCCAGCATTCGGTTTGGCTGATGGCTGCTGGCGAACGGCAAC 1870	Db	284	----- 284	
Qy	1931	TTTGATCTGGGAAAACCTGGCCATGCTAAAGGNGGGCTGATGGCACCAACTGGAA 1990	Qy	1871	AAATCCGGTGCACGGGGACCGSTACTGACATGCTGCTGCTGGCTGCTGGCGAAC 1930	
Db	290	----- 290	Db	285	----- 285	
Qy	1991	ATCGACACAAAGTGGTCGGGTGACATAAGCTGGCTGATGAACTGGCTGAA 2050	Qy	1991	ATCGACACAAAGTGGTCGGGTGACATAAGCTGGCTGATGAACTGGCTGAA 2050	

Copyright GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6		Copyright (c) 1993 - 2005 Compugen Ltd.		Description	
Result No.	Score	Query Match	Length	DB ID	
OM nucleic - protein search, using frame_plus_n2p model					Sequence 1, Appli
Run on:	August 2, 2005, 18:23:19 ; Search time 801 Seconds (without alignments)				Sequence 2, Appli
Scoring table:	BLOSUM62				Sequence 3, Appli
Scoring table:	Xgapext 0.0 , Ygapext 0.5 Ygapop 10.0 , Xgapext 0.5 Ygapop 10.0 , Fgapext 7.0 Fgapop 6.0 , Delext 7.0				Sequence 4, Appli
Searched:	1745140 seqs, 389608008 residues				Sequence 5, Appli
Total number of hits satisfying chosen parameters:	3490280				Sequence 6, Appli
Minimum DB seq length: 0					Sequence 7, Appli
Maximum DB seq length: 2000000000					Sequence 8, Appli
Post-processing: Minimum Match 0%					Sequence 9, Appli
Post-processing: Maximum Match 100%					Sequence 10, Appli
Listing first 45 summaries					Sequence 11, Appli
Command line parameters:					Sequence 12, Appli
-MDBL=frame_n2p model -DEV=x1P -Q=/cgn2_1/USPFO_spool/p/US10649433/runat_02082005_164946_1571/app_query.fasta_1.6919 -QB=public_b1/USPFO_spool/p/US10649433/runat_02082005_164946_1571/app_query.fasta_1.6919 -TRANS=human40_cdi -LIST=45 -DOCALIGN=200 -THR SCORE=0CT -THR MAXX=100 -MINLEN=0 -MAXLEN=0 -ALIGN=15 -MODB=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIS200 -USER=US10649433 @CGN 1.1 964 @runat_02082005_164946_1571 -NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=30 -WARN TIMEOUT=120 -THREADES=1 -XGAPPOP=10 -XGAPEXT=0.5 -FGAPPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7					Sequence 13, Appli
Published Applications AA:*					Sequence 14, Appli
1: /cgn2_6/podata/1/pubpaas/US07_PUBCOMB.pep:*					Sequence 15, Appli
2: /cgn2_6/podata/1/pubpaas/PCT_NEW_PUB.pep:*					Sequence 16, Appli
3: /cgn2_6/podata/1/pubpaas/US05_NEW_PUB.pep:*					Sequence 17, Appli
4: /cgn2_6/podata/1/pubpaas/US07_NEW_PUB.pep:*					Sequence 18, Appli
5: /cgn2_6/podata/1/pubpaas/PCT05_PUBCOMB.pep:*					Sequence 19, Appli
6: /cgn2_6/podata/1/pubpaas/PCT05_PUBCOMB.pep:*					Sequence 20, Appli

PREDICTION

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT	1
US-10477-044-1	
SEQUENCE 1, Application US/10477044	
; Publication No. US2004025902A1	
GENERAL INFORMATION:	
; APPLICANT: Fricker, Mark	
; TITLE OF INVENTION: Universal Fluorescent Se	
; FILE REFERENCE: HO-P0295US0	
; CURRENT APPLICATION NUMBER: US/10/477,044	
; CURRENT FILING DATE: 2003-11-07	
; PRIOR APPLICATION NUMBER: PCT/GBO/02/02183	
; PRIOR FILING DATE: 2002-02-10	
; PRIOR APPLICATION NUMBER: GB 0111459.4	
; PRIOR FILING DATE: 2001-05-10	
; NUMBER OF SEQ ID NOS: 12	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 1	
; LENGTH: 1967	
; TYPE: PRT	

SUMMARIES

7: /cgn2_6_ptodata/1/pubpaa/us08_new_pub.pep:*

8: /cgn2_6_ptodata/1/pubpaa/_pubcomb.pep:*

9: /cgn2_6_ptodata/1/pubpaa/us09_pubcomb.pep:*

10: /cgn2_6_ptodata/1/pubpaa/us09b_pubcomb.pep:*

11: /cgn2_6_ptodata/1/pubpaa/us09c_pubcomb.pep:*

12: /cgn2_6_ptodata/1/pubpaa/us09_new_pub.pep:*

13: /cgn2_6_ptodata/1/pubpaa/us10a_pubcomb.pep:*

14: /cgn2_6_ptodata/1/pubpaa/us10b_pubcomb.pep:*

15: /cgn2_6_ptodata/1/pubpaa/us10c_pubcomb.pep:*

16: /cgn2_6_ptodata/1/pubpaa/us10d_pubcomb.pep:*

17: /cgn2_6_pcodata/1/pubpaa/us10e_pubcomb.pep:*

18: /cgn2_6_ptodata/1/pubpaa/us10_new_pub.pep:*

19: /cgn2_6_ptodata/1/pubpaa/us11a_pubcomb.pep:*

20: /cgn2_6_ptodata/1/pubpaa/us11_new_pub.pep:*

21: /cgn2_6_ptodata/1/pubpaa/us60_new_pub.pep:*

22: /cgn2_6_ptodata/1/pubpaa/us60__pubcomb.pep:*

Result No.	Score	Query Match	Length	DB ID	Description
1	4344.4	64.6	5271	12 CVU899316	U899316 Cloning vec
2	3855.3	57.3	4806	12 AF05044	AF05044 Expressio
3	3854.2	57.3	4205	6 AX576406	AX576406 Sequence
4	3853.4	57.3	5013	6 AR264513	AR264513 Sequence
5	3853.2	57.3	4203	6 AX576404	AX576404 Sequence
6	3853.2	57.3	4204	6 AX576405	AX576405 Sequence
7	3853.2	57.3	4217	6 AX576408	AX576408 Sequence
8	3853.2	57.3	4218	6 AX576409	AX576409 Sequence
9	3852.8	57.3	4216	6 AX576407	AX576407 Sequence
10	3852.2	57.2	4176	6 A13038	A13038 Artificial
11	3852.2	57.2	4176	12 M22744	M22744 Cloning vec
12	3852.2	57.2	4239	6 A59289	A29289 pSBE-Bpi co
13	3851.6	57.2	4476	6 AR102989	AR102989 Sequence
14	3851.6	57.2	4476	6 AR172783	AR172783 Sequence
15	3851.4	57.2	4214	6 BD273391	BD273391 P40racet
16	3851.4	57.2	4214	6 AX027488	AX027488 Sequence
17	3850.6	57.2	4176	12 XXU13872	XXU13872 pTrc99a Clo
18	3824.8	56.8	4232	12 SYNLACIQ	L24193 Cloning vec
19	3824.8	56.8	4735	12 SYNBTALAC	U14367 Cloning vec

			Aad45059	Transposo
			Adg4681	Donor pl
			Aax78874	Human tis
			Aax78896	Human tis
			Aax78883	Human tis
			Aax78893	Human tis
			Aax78893	Human tis
			Aaa59260	Nucleotid
			Aaa62632	Vector PE
			Aan90709	Sequence
			Adi36476	Plasmid C
			Adg46819	PEM DNA
			Adg14707	Mouse OB
			Aao05397	Secretion
			Adh45226	Modified
			Acf0051	Plasmid P
			Adl18668	Plasmid P
			Auc55454	Destinat
			Ab258765	Destinat
			Aan91062	Sequence
			Adf50156	Plasmid C
			Aac55467	Destinat
			Aac55464	Destinat
			Adz41729	Plasmid P
			Aat90491	Vector pl
C	21	3852.2	57.2	6
C	22	3852.2	57.2	5926
C	23	3851.8	57.2	5926
C	24	3851.8	57.2	5099
C	25	3851.8	57.2	5132
C	26	3851.8	57.2	5324
C	27	3851.8	57.2	5342
C	28	3851.4	57.2	5342
C	29	3851.4	57.2	4214
C	30	3850.6	57.2	4176
C	31	3850.6	57.2	12
C	32	3850.6	57.2	ADG46818
C	33	3851.4	57.1	5103
C	34	3727.8	55.4	1489
C	35	3542	52.6	5161
C	36	3344.8	49.7	10
C	37	3344.8	49.7	ADH45226
C	38	3341.2	49.7	9
C	39	3341.2	49.7	ACF0051
C	40	3241.8	48.2	12
C	41	3230.2	48.0	ADL18668
C	42	2864.8	42.6	3
C	43	2864.8	42.6	AUC55454
C	44	2863.2	42.6	9
	45	2858	42.5	4557
Copyright	Gencore version 5.1.6			
(c) 1993 - 2005	Compugen Ltd.			
DM nucleic - nucleic search, using sw model				
run on:	August 4, 2005, 19:20:28 ; Search time 3192 Seconds			
	(without alignments)			
	12479.304 Million cell updates/sec			
Title:	US-10-649-433C-1			
Perfect score:	6729			
Sequence:	1 gtttgcacacttcatatcg.....agttagcgcgaattgtatctg 6729			
Scoring table:	IDENTITY_NUC			
	Gapext 1.0 , Gapext 1.0			
searched:	4390206 seqs, 2959870667 residues			
Total number of hits satisfying chosen parameters:	8780412			
Minimum DB seq length:	0			
Maximum DB seq length:	2000000000			
Post-processing: Minimum March 0%				

ALIGNMENTS

RESULT 1	ABZ23939	standard; DNA; 5903 BP.
	ID ABZ23939	
	XX	
	AC AC	
	XX DT	
	18-MAR-2003 (first entry)	
	XX XX	
	DE XX	
	KW Fluorescent; fluorescence resonance energy transfer; FRET; tumour; estrogen; tamoxifen; therapy; pTRCCFRET3 ; ds.	
	XX OS Synthetic.	
geneseqm1980bs:	*	
geneseqm1990bs:	*	
geneseqm2000bs:	*	
geneseqm2001as:	*	
geneseqm2001bs:	*	
geneseqm2002as:	*	
geneseqm2002bs:	*	
geneseqm2003as:	*	
geneseqm2003bs:	*	
geneseqm2003cs:	*	
geneseqm2003ds:	*	
geneseqm2004as:	*	
geneseqm2004bs:	*	
geneseqm2004cs:	*	

SUMMARIES

result	No.	Query			DB	ID	Description
		Score	Match	Length			
	1	4119.8	61.2	5903	8	AB223939	Nucleotid
	2	3855	57.3	5027	2	AAX78865	Aax78865 Human tis
	3	3855	57.3	5069	2	AAX78871	Human tis
	4	3854.5	57.3	4205	6	ABQ74927	Avidity I
	5	3853.4	57.3	4816	4	AA007799	pTRCHisB
	6	3853.4	57.3	4816	12	Adf72546	Mouse Adi
	7	3853.4	57.3	4816	12	Adg14708	Globular
	8	3853.4	57.3	5013	9	AcE63355	E. coli D
	9	3853.4	57.3	5013	10	Ad59423	E. coli pu
	10	3853.4	57.3	5013	10	AdE86127	E. coli P
	11	3853.4	57.3	5104	4	Adf72545	pTRCHisB
	12	3853.4	57.3	5104	12	Adf72545	Mouse Adi
	13	3853.2	57.3	4203	6	ABQ74925	Avidity I
	14	3853.2	57.3	4204	6	ABQ74926	Avidity I
	15	3853.2	57.3	4217	6	ABQ74929	Avidity I
	16	3853.2	57.3	4218	6	ABQ74930	Avidity I
	17	3852.8	57.3	4216	6	ABQ74928	Avidity I
C	18	3852.2	57.2	5926	2	Aav32978	Tn7 donor
C	19	3852.2	57.2	5926	2	Aav32977	Tn7 donor
C	20	3852.2	57.2	5926	6	Ad45050	DM del tra

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

"codons"

CDS 2 .5912 /*tag= a /note= "contains internal stop codons"
W0200290987-A2.

Probe, useful e.g. for medical diagnosis, and detection of pollutants in water systems and contaminants in foodstuffs, has target binding site moiety and mimic moiety attached to two fluorescent polyptides, and linker.

moiety attached to a first fluorescent polypeptide, a mimic moiety capable of binding to the target binding site moiety and attached to a second fluorescent polypeptide, and a linker connecting the two fluorescent polypeptides and which allows the distance between the fluorescent polypeptides to vary. (I), or a cell harbouring (I), or a polynucleotide which encodes (I), or a sensor (IV) comprising (I), a light source which is capable of exciting the probe, and a detecting which is capable of measuring the amount of fluorescence resonance energy transfer (FRET) from the probe for detecting the presence or absence of a target substance in a test sample, and for identifying an inhibitor of binding between two substances, where the two substances would bind to each other in the absence of an inhibitor. (I) is useful for medical diagnosis, detection of pollutants in water systems, and detection of contaminants in foodstuffs and in animal and plant biology and for identification of new therapeutic substances. (I) is useful to screen for stimulators, which increase or promote binding between two substances, to identify a factor which increases the strength of binding between two substances or a factor whose presence is necessary for the binding of two substances to take place, to identify stimulators and/or stabilizers of binding interactions, to detect the presence of a substance, for e.g. metabolite, hormone, drug, or pollutant in an extract, for e.g. a fluid sample derived from any organism, including an animal or human, plant, fungus or microbe. (I) is also useful for detecting sugars, oligoaccharides or non-carbohydrate mimetics, to determine the presence or absence of steroid hormones, for diagnosis, drug and alcohol testing and testing for exposure to toxins or pollutants, to detect air-borne substances for e.g. atmospheric pollutants if these substances are soluble, and to detect specific substances in plant, fungal or microbial extracts. Plant extracts for e.g. exudates are useful in determining the presence of plant pathogenic viruses or bacteria in plant. Additionally, the probes may be used to determine the presence and amount of trace elements or pollutants in plant extracts. Thus results of such assays may provide indirect measurements of soil quality and as indicative of particular types of soil pollution. Further application of the probes is to detect proteins expressed in transgenic plants or transgenic animals, fungi or microbes, and to detect contaminants or pollutants in water supplies, soil or factory effluents, and in quality control for detecting substances, for e.g. contaminants in foodstuffs and medicaments. The probes are also useful for screening for inhibitors of growth of estrogen-sensitive breast tumours, which is useful for identifying anti-tumour agents that act at a site distinct from that targeted by the synthetic estrogen, tamoxifen. The probe is also useful for identifying protease inhibitors, screening for active binding site inhibitors of the protease, and for identifying intracellular G protein signal inhibitors. The inhibitors identified by the probe are useful for treating human or animal body by therapy. The present sequence represents the nucleotide sequence of plasmid pTRCFRET3

Sequence 5903 BP; 1429 A; 1594 C; 1608 G; 1272 T; 0 U; 0 Other;

Qy	2369	GATGACGGGAACCTAACAGACCGTGTCTGAAGTCAGTTGAAGTGATAACCCCTGTTAA	2428
Db	1205	GACGACGGCAACTAACAGACCCGGAGGTGAAGTTGAGCTGGACACCCCTGGTGAAC	1264
Qy	2429	CGTATGAGTTAAAGGTATTGATTTAAGAAGTGGAAACATTCGACAACTC	2488
Db	1265	CGCATGAGTGAAAGGCATCGACTCAAGGACGGGACATCGACTCTGGGCACAGCTG	1324
Qy	2489	GAGTACAACCTAACTCACACAATTTATACATCAGGGAGACAAACAAAATGGAATC	2548
Db	1325	GAGTACAACCTAACAGACCCATATCGGCCAACACAGAAACGCCAT	1384
Qy	2549	AAGGCTAACCTCAAAATTGCCAACACATGAAAGTGGATCCGTTCAACTAGCAGACCAT	2608
Db	1385	AAGGTGAACTTCAGATCCCAACTGAGGCGGAGCGTGCACGTCGCCACCAC	1444
Qy	2609	TATCACAAATACTCCAATTGGCAATGGCCCTGTCCTTTACGACACCAATTACCTG	2668
Db	1445	TACCGGAGAACCCCCATGGCACGGGCGCTGCTGTGCCGACACCACTACCTG	1504
Qy	2669	TCGACACAAATCTGCCCTTGGAAAGAACATCCAAACGGTACACATGGTCTCTCRT	2728
Db	1505	AGCTACCGTCGCCCTGAGCAAAGAACCCAACTCGGGATCTACATGGTCTCGTGT	1564
Qy	2729	GAGTTGTACTGCTGGATTAACATGGATGGAGGTCTACAAATAAAAGCTT	2788
Db	1565	GAGTTGTAACGCCGGGATCTACTCTGGCA-----	1598
Qy	2789	ACGTAAACAAAAACTCATCTAGAGGATCTGAATAGGCCCTGACCATCATC	2848
Db	1599	-----	1598
Qy	2849	ATCATCATGGTTAACCGGTCTACGTTGGCTTTGGGATGAGAAAGATT	2908
Db	1599	-----	1599
Qy	2909	CAGCCGATACAGATTAAATCAGAACGCAAGCGCTGATAAAACAGATTGCTGG	2968
Db	1633	CAGCTGTATACAGATTAAATCAGACGCAAGCGCTGATAAAACAGATTGCTGG	1692
Qy	2969	CGGCAGTAGCCGGGGTCCACCTGACCCATGCCGAACCTCAGAAGTGAACGGCTAG	3028
Db	1693	CGGCAGTAGCCGGGGTCCACCTGACCCATGCCGAACCTCAGAAGTGAACGGCTAG	1752
Qy	3029	CGCCATGGTAGTGTGGGTCTCCCTGAGCTGGGAGTAGGAAACTCCGGGAT	3088
Db	1753	CGCCATGGTAGTGTGGGTCTCCCTGAGCTGGGAGTAGGAAACTCCGGGAT	1812
Qy	3089	AACGAAGGGTCAGCTGAAGAACGCTGGCCTTTCGTTATCTGTTGTTGCTGGTAACG	3148
Db	1813	AACGAAGGGTCAGCTGAAGAACGCTGGCCTTTCGTTATCTGTTGTTGCTGGTAACG	1872
Qy	3149	CTCTCTGAGTAGGCAAACTCGCCGGAGGGATTGAGCTTGCGTCAAGGAAACGCCCG	3208
Db	1873	CTCTCTGAGTAGGCAAACTCGCCGGAGGGATTGAGCTTGCGTCAAGGAAACGCCCG	1932
Qy	3209	GAGGTGGCGCGAGACGCCATTAAGGAAAGGCC	3268
Db	1933	GAGGTGGCGCGAGACGCCATTAAGGAAAGGCC	1992
Qy	3269	TCCTGACGGATGGCCTTTCGCTTCTACAAACTCTTGTATTCTAAATACA	3328
Db	1993	TCTGAGGGATGGCTTCTACAAACTCTTGTATTCTAAATACA	2052
Qy	3329	TTCAAATATGATCGCTCATGAGACAATAACCTGATAATATGAA	3388
Db	2053	TTCAAATATGATCGCTCATGAGACAATAACCTGATAATATGAA	2112
Qy	3389	AGGAGAGPATGAGTGTGAACTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTG	3448
Db	2113	AGGAGAGPATGAGTGTGAACTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTG	2177

1145 GACTCTTCAGTCGGCATGCCGAGGTACGTCAAGGCAACATCTCTCAAG

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW mode!

Run on: August 4, 2005, 09:18:00 : Search time 992 Seconds
(without alignments)

11059.300 Million cell updates/sec

Title: US-10-649-433C-1

Perfect score: 6729

Sequence:

Scoring table: IDENTITY NUC Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138159 residues

Total number of hits satisfying chosen parameters: 2405558

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing First 45 summaries

Database : Issued_Patents NA:
1: /cgn2_6/_ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/_ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/_ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/_ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/_ptctus/1/ina/ptctus_COMB.seq:
6: /cgn2_6/_ptodata/1/ina/backtiles.seq:
*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	3853.4	57.3	5013	4	US-09-183-188B-5		Sequence 5, Appli
c 2	3853.2	57.2	5926	3	US-09-07-169-3		Sequence 3, Appli
c 3	3853.2	57.2	5926	3	US-09-037-169-4		Sequence 4, Appli
c 4	3853.6	57.2	4476	3	US-08-344-2		Sequence 2, Appli
5	3853.6	57.2	4476	3	US-09-498-599-2		Sequence 2, Appli
6	2729	40.6	4557	4	US-08-778-717-5		Sequence 5, Appli
c 7	2598.8	38.6	4357	4	US-09-955-872-10		Sequence 10, Appli
c 8	2598.8	38.6	4533	3	US-08-801-344-1		Sequence 1, Appli
c 9	2598.8	38.6	4593	3	US-09-498-599-1		Sequence 1, Appli
10	2596	38.6	6312	1	US-08-531-601-3		Sequence 3, Appli
11	2596	38.6	6312	2	US-08-853-032-3		Sequence 3, Appli
12	2591.8	38.5	5238	6	5453363-1		Patent No. 5453363
13	2591.8	38.5	5238	6	5453363-1		Sequence 1, Appli
c 14	2483.6	36.9	4410	1	US-08-534-469-1		Sequence 1, Appli
c 15	2483.6	36.9	4410	2	US-08-906-957-1		Sequence 1, Appli
c 16	2489.8	36.3	5446	4	US-09-358-858C-1		Sequence 1, Appli
c 17	2188.4	32.5	7892	2	US-07-946-09A-40		Sequence 40, Appli
c 18	2155.6	32.5	7633	3	US-09-028-851-1		Sequence 1, Appli
c 19	2165.6	32.5	7633	3	US-08-815-520-1		Sequence 1, Appli
c 20	2185.6	32.5	7633	3	US-09-273-163-1		Sequence 1, Appli
c 21	2183.4	32.4	5825	4	US-08-809-513A-7		Sequence 7, Appli
c 22	2183.4	32.4	8068	3	US-09-301-593-27		Sequence 27, Appli
c 23	2183.4	32.4	8068	3	US-09-301-593-35		Sequence 35, Appli
c 24	2182.6	32.4	9144	3	US-08-556-978B-79		Sequence 79, Appli
c 25	2182.6	32.4	7731	3	US-09-301-593-29		Sequence 29, Appli
c 26	2182.6	32.4	7731	3	US-09-301-593-42		Sequence 42, Appli
c 27	2182	32.4	3699	1	US-08-053-131-120		Sequence 120, App

ALIGNMENTS

RESULT 1
US-09-183-188B-5
; Sequence 5, Application US/09183188B
; Patent No. 6431905

; GENERAL INFORMATION:
; APPLICANT: Sorscher, Eric J.
; APPLICANT: Parker, William B.
; APPLICANT: Vaud, William B.
; APPLICANT: Gadi, Vijaykrishna K.
; TITLE OF INVENTION: RECOMBINANT E. COLI FOR DELIVERY OF PNP TO TUMOR CELLS
; FILE REFERENCE: UAB-12405/22
; CURRENT FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 08/193,188B
; PRIOR FILING DATE: 1993-09-14
; PRIOR APPLICATION NUMBER: 60/064,676
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 5
; LENGTH: 5013
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-183-188B-5

Query Match 57.3%; Score 3853.4%; DB 4; Length 5013;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3854; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2875 AGCTGGACTGTTGGGGATGAGAGATTTCAGCTGATAAGATTAAATCAGAAC 2934
Db 1159 AGCTGGCTGTTGGGGATGAGAGATTTCAGCTGATAAGATTAAATCAGAAC 1218

Query Match 57.3%; Score 3853.4%; DB 4; Length 5013;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3854; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2935 GCAGAAGGGCTGTATAAACAGATTGCGCTGGCCGAGTAGGCCGGTGGTCCACCTG 2994
Db 1219 GCAGAAGGGCTGTATAAACAGATTGCGCTGGCCGAGTAGGCCGGTGGTCCACCTG 1278

Query Match 57.3%; Score 3853.4%; DB 4; Length 5013;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3854; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2995 ACCCATGGCAACTCAGAACTGAAAGCTGCGGATGTAATGTCGGCTTCCCC 3054
Db 1279 ACCCATGGCAACTCAGAACTGCGGATGTAATGTCGGCTTCCCC 1338

Query Match 57.3%; Score 3853.4%; DB 4; Length 5013;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3854; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3055 ATGCAGAGATGGGAATGGCAACTGCGGATGTAATGTCGGCTTCCCC 3114
Db 1339 ATGCAGAGATGGGAATGGCAACTGCGGATGTAATGTCGGCTTCCCC 1398

Query Match 57.3%; Score 3853.4%; DB 4; Length 5013;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3854; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3115 GCCTTTCGTTTTATCTGTTGTCGTTGACGCTCTCTGAGTAGGACAATCCCG 3174
Db 1399 GCCTTTCGTTTTATCTGTTGTCGTTGACGCTCTCCGAGACAATCCCG 1458

Qy	3175	GGAGCCGATTTCGAACCGTTGCCGAAAGCCACGGCCCCGGTGGGGCAGGACGCCCA	3234	4255	GGTAACTGTGCGACCAAGTTACTCATATACTTTAGATTGATTAAACTTCACTTTT	4314	
Db	1459	GGAGCGGATTTCGAACCTTGCGGAAAGCCACGGGGGGTGGGGCAGGACGCCCA	1518	Db	2539	GGTAACTGTGCGACCAAGTTACTCATATACTTTAGATTGATTAAACTTCACTTTT	2598
Qy	3235	TAAACTGCAAGCCATCAAATAAGCGAAGGCCATCTGA CGGATGCGCTTTTGGTT	3294	Qy	4315	AATTAAAGGATCPAGGTGAGAACCTTCTGATAATCTGACCCAAAATCCCTTAAC	4374
Db	1519	TAAACTGCAAGCCATCAAATAAGCGAAGGCCATCTGA CGGATGCGCTTTTGGTT	1578	Db	2599	AATTAAAGGATCPAGGTGAGAACCTTCTGATAATCTGACCCAAAATCCCTTAAC	2658
Qy	3295	CTACAACTCTTTTGTATTTTCTAAATACATTGATCCTGCAGAGAC	3354	Qy	4375	GTGAGTTTCTCACTGAGCTGAGACCCCAGTAAAGATCAAAGGATCTCTGTAG	4434
Db	1579	CTACAACTCTTTTGTATTTTCTAAATACATTGATCCTGCAGAGAC	1638	Db	2659	GTGAGTTTCTCACTGAGCTGAGACCCCAGTAAAGATCAAAGGATCTCTGTAG	2718
Qy	3355	AATAACCCGATAAAGCTCCATATAATTTGAAAGGAGAGTAGTGAAGATTACATT	3414	Qy	4435	ATCCCTTTTCTGCCGTAATCTGCTGCAACAAAAACCCACCGTACAGCGG	4494
Db	1639	AATAACCCGATAAATGCTTCAATAATTTGAAAGGAGAGTAGTGAATTCACATT	1698	Db	2719	ATCCCTTTTCTGCCGTAATCTGCTGCAACAAAAACCCACCGTACAGCGG	2778
Qy	3415	TCCGTTGCGCTTAAITCCCTTTGCGGATTTGCTCACCCAG	3474	Qy	4495	TGGTTGTTTCCGGATCAAGGCTTACCAACTCTTTTCCGAAAGGTAACGGCTCAGCA	4554
Db	1699	TCCGTTGCGCTTAAITCCCTTTGCGGATTTGCTCACCCAG	1758	Db	2779	TGGTTGTTTCCGGATCAAGGCTTACCAACTCTTTTCCGAAAGGTAACGGCTCAGCA	2838
Qy	3475	AAACGCTGTGAAAGTAAAGATGCTGAAGATCACTGGTGCACCAAGTGGTTACATCG	3534	Qy	4555	GAGCGAGATAACCAAAATCTGCTCTGCTAATCCCTTTACCGATGGCTGTGCCA	4614
Db	1759	AAACGCTGTGAAAGTAAAGATGCTGAAGATCACTGGTGCACCAAGTGGTTACATCG	1818	Db	2839	GAGCGAGATAACCAAAATCTGCTCTGCTAATCCCTTTACCGATGGCTGTGCCA	2898
Qy	3535	AACTGGSATCTCAACASCGGTAAGATCTGAGAGTTTCGCCGAGAACGTTTICCAA	3594	Qy	4615	ACTCTTAGACCCGCTTACATCTGCTCTGCTAATCCCTTTACCGATGGCTGTGCCA	4674
Db	1819	AACTGGATCTCAACAGGGTAGAGATCTTGTAGATCTTGTAGTTTCCAA	1878	Db	2899	ACTCTTAGACCCGCTTACATCTGCTCTGCTAATCCCTTTACCGATGGCTGTGCCA	2958
Qy	3595	TGATGAGCACTTTAAGTTGCTATGTTGCTATGTTGCGGGTATATCCCTGTTGACCGGGC	3654	Qy	4675	GTGGCGATAATGCTGTTCTACCGGTTGGACTCAAGACCTGTTACCGGATAAGGCGC	4734
Db	1879	TGATGAGCACTTTAAGTTGCTATGTTGCGGGTATATCCCTGTTGCTACGGGGC	1938	Db	2959	GTGGCGATAATGCTGTTCTACCGGTTGGACTCAAGACCTGTTACCGGATAAGGCGC	3018
Qy	3655	AGAGGCAACTCTGGCGATAACATCTGCTATGTTGCGGGTATATCCCTGTTGCTACGGGGC	3714	Qy	4735	ACCGTTCGGCTGAAAGGGGGTTCTGGCACACCCACCTGGAGCAACGACCTACA	4794
Db	1939	AGAGGCAACTCTGGCGATAACATCTGCTATGTTGCGGGTATATCCCTGTTGCTACGGGGC	1998	Db	3019	ACGGCTTCGGCTGAAAGGGGGTTCTGGCACACCCACCTGGAGCAACGACCTACA	3078
Qy	3715	TCAAGAAAAGCAACTCTAACGATGCTGATGAGCTTACGAGTTACGAGTAAAGGAA	3774	Qy	4795	CGGAATCTGAGATACTCTAGGGTGTACCTACGGGTGAGCTATGAAAGGCTA	4854
Db	1999	TCAAGAAAAGCAACTCTAACGATGCTGATGAGCTTACGAGTAAAGGAA	2058	Db	3079	CGGAATCTGAGATACTCTAGGGTGTACCTACGGGTGAGCTATGAAAGGCTA	3138
Qy	3775	CCATGAGTGTAAACACTGGCCAACCTACTCTGCAACCATGGAGCCAGGG	3834	Qy	4855	AGGGCGACAGGTATCGGTAAAGCGGCGAGGGTGGAAAGGAGCCACGGCTTC	4914
Db	2059	CCATGAGTGTAAACCTACTCTGCAACCATGGAGCCAGGG	2118	Db	3139	AGGGCGACAGGTATCGGTAAAGCGGCGAGGGTGGAAAGGAGCCACGGCTTC	3198
Qy	3835	TAACCGCTTTTGCAACATGGGGCATCTGTAACCTGCTGTTGACCG	3894	Qy	4915	GCGGGCAAAACCCCTGGTACTTCTGTTGGGTTTTCGCACTCTGACTGTGCC	4974
Db	2119	TAACCGCTTTTGCAACATGGGGCATCTGTAACCTGCTGTTGACCG	2178	Db	3199	CAGGGCGAAAGCCCTGGTATCTTATAGCTCTGTTGGGTTTCGCACTCTGACTGTGCC	3258
Qy	3895	AGCTGTAAATGAAAGCCATACAAACGAGCGCTGAACTGGAGCCAA	3954	Qy	4975	GTCGATTTTGTGATGCTGTTGGGAGGCTTATGAAAAACGCCAGCAAGGGG	5034
Db	2179	AGCTGTAAATGAAAGCCATACAAACGAGCGCTGAACTGGAGCCAA	2238	Db	3259	GTCGATTTTGTGATGCTGTTGGGAGGCTTATGAAAAACGCCAGCAAGGGG	3318
Qy	3955	CAACGTTGCGCAAACCTACTGCGCAACTTACTCTGCAACCTGGGCAACATTAA	4014	Qy	5035	CCCTTTACGGTTCTGGCTTTCGCTACATGGTTCTGGCTTACCTGGTTATGGCTT	5094
Db	2239	CAACGTTGCGCAAACCTACTGCGCAACTTACTCTGCAACCTGGGCAACATTAA	2298	Db	3319	CCCTTTACGGTTCTGGCTTTCGCTACATGGTTCTGGCTTACCTGGTTATGGCTT	3378
Qy	4075	GCTGGTTTATCTGTATAATCTGACCCGTTGAGCTGAGCTGAGCTGAG	4134	Qy	5095	CCCTTTACGGTTCTGGCTTTCGCTACATGGTTCTGGCTTACCTGGTTATGGCTT	5154
Db	2359	GCTGGTTTATCTGTATAATCTGACCCGTTGAGCTGAGCTGAGCTGAG	2418	Qy	3379	CCCTTTACGGTTCTGGCTTTCGCTACATGGTTCTGGCTTACCTGGTTATGGCTT	3438
Qy	4135	CACTGGGGCCAGATGGTAAAGCCTCCGTTATGCTGTTGGGAGCTGG	4194	Qy	5155	GCCGAGCAGCCAGGCGAGCCAGGCTGAGCTGAGCTGAGCTGAGCTGAG	5214
Db	2419	CACTGGGGCCAGATGGTAAAGCCTCCGTTATGCTGTTGGGAGCTGG	2478	Db	3439	GCCGAGCAGCCAGGCGAGCCAGGCTGAGCTGAGCTGAGCTGAGCTGAG	3498
Qy	4195	CAACTATGATGAAACGAAATAAGACAGATGGTGAAGCTGAGTAAAGGATT	4254	Qy	5215	ATTTCCTCCTTAACGGTATCTGGCTGAGCTGAGCTGAGCTGAGCTGAG	5274
Db	2479	CAACTATGATGAAACGAAATAAGACAGATGGTGAAGCTGAGTAAAGGATT	2538	Db	3499	ATTTCCTCCTTAACGGTATCTGGCTGAGCTGAGCTGAGCTGAGCTGAG	3558
Qy				5325	TCTGCTGATGAGCCGATAGTTAACGGCAACTCTGGCTATGCTACTGGT	5334	
				5359	TCTGCTGATGAGCCGATAGTTAACGGCAACTCTGGCTATGCTACTGGT	3618	
				5335	CATGGCTGCGCCGACACCCGGCAACCCGGCAACCCGGCAACCCGGCAAC	5394	

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2005, 15:42:19 ; Search time 3779 Seconds

(without alignments)
11542.622 Million cell updates/sec

Title: US-10-649-433C-1.

Perfect score: 6729

Sequence: 1 gtttagacgttcatcgat.....agttagcgcgaattgtatcg 6729

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters:

14594722

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/_ptodata/1/pubnra/US07_PUBCOMB.seq:*

2: /cgn2_6/_ptodata/1/_pubnra/PCT_NEW_PUB.seq:*

3: /cgn2_6/_ptodata/1/_pubnra/US06_NBW_PUB.seq:*

4: /cgn2_6/_ptodata/1/_pubnra/US07_PUBCOMB.seq:*

5: /cgn2_6/_ptodata/1/_pubnra/US07_NEW_PUB.seq:*

6: /cgn2_6/_ptodata/1/_pubnra/BCTUS_PUBCOMB.seq:*

7: /cgn2_6/_ptodata/1/_pubnra/US08_NEW_PUB.seq:*

8: /cgn2_6/_ptodata/1/_pubnra/US08_PUBCOMB.seq:*

9: /cgn2_6/_ptodata/1/_pubnra/US09_PUBCOMB.seq:*

10: /cgn2_6/_ptodata/1/_pubnra/US09_PUBCOMB.seq:*

11: /cgn2_6/_ptodata/1/_pubnra/US09_PUBCOMB.seq:*

12: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

13: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

14: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

15: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

16: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

17: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

18: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

19: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

20: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

21: /cgn2_6/_ptodata/1/_pubnra/US10_PUBCOMB.seq:*

22: /cgn2_6/_ptodata/1/_pubnra/US10_NEW_PUB.seq:*

23: /cgn2_6/_ptodata/1/_pubnra/US1A_PUBCOMB.seq:*

24: /cgn2_6/_ptodata/1/_pubnra/US60_NEW_PUB.seq:*

25: /cgn2_6/_ptodata/1/_pubnra/US60_PUBCOMB.seq:*

26: /cgn2_6/_ptodata/1/_pubnra/US60_PUBCOMB.seq:*

RESULT 1
US-10-649-433C-1 ; Sequence 1, Application US/10649433C ; Publication No. US/0050118726A1 ; GENERAL INFORMATION: ; APPLICANT: SCHULITZ, Jerome S. ; TITLE OF INVENTION: System and Method for Detecting Bioanalytes and Method for Producing a Bioanalyte Sensor ; FILE REFERENCE: 03-016 ; CURRENT APPLICATION NUMBER: US/10/649,433C ; CURRENT FILING DATE: 2003-08-26 ; NUMBER OF SEQ ID NOS: 1 ; SOFTWARE: Patentin version 3.3 ; SEQ ID NO: 1 ; LENGTH: 6729 ; TYPE: DNA ; ORGANISM: Escherichia coli ; US-10-649-433C-1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Query Match	Length	DB ID	Description
1	6729	100.0	6729	21	US-10-649-433C-1
2	3853.4	57.3	5013	14	US-10-198-034-5
3	3853.4	57.3	5013	15	US-10-198-288-5
c	3852.2	57.2	5926	13	US-10-024-809-3
c	3850.6	57.2	4176	16	US-10-244-424A-3
c	3344.8	49.7	5024	17	US-10-313-963A-52

Query Match	Best Local Similarity	Score	DB 21;	Length	6729;
Qy	1	GTTGACAGCTTATCATGACTGCAGGTCAATGCCTCTGGCTCAGGCATC	60		
Db	1	GTTGACAGCTTATCATGACTGCAGGTCAATGCCTCTGGCTCAGGCATC	60		
Qy	61	GGAGCTGTGTGTTAGCTGTGTCAGGTGTAATACTACTGCATAATTCTGTGCTCAAGGC	120		
Db	61	GGAGCTGTGTGTTAGCTGTGTCAGGTGTAATACTGCATAATTCTGTGCTCAAGGC	120		

Qy	121	GGACTCCGTTCTGGATAATTTTTCGGCGACATCATAACGGTTCTGCCAATTATTC	180	Db	1201	GGCTATTGAGCAAGATGCGAAAAGCCGCCAGATCTTCAGTCGTGATGAAATGTTCTCA	1260
Db	121	GGACTCCGTTCTGGATAATTTTTCGGCGACATCATAACGGTTCTGCCAATTATTC	180	Qy	1261	GAATGACCAGTCAGCCAGAACGATCAGATCGACCTATTCCTGCCAAGGGGTGAGGC	1320
Qy	181	TGAAATCA GCCTTGACAATTAACTCATC CGCTCGTATAMTGTGTAAGTGTGCGA	240	Db	1261	GAATGACCAGTCAGCCAGAACGATCAGATCGACCTATTCCTGCCAAGGGGTGAGGC	1320
Db	181	TGAAATCA GCCTTGACAATTAACTCATC CGCTCGTATAMTGTGTAAGTGTGCGA	240	Qy	1321	ACTGGCCATCACCTCGTGAACCCCGACGTGCGCTACCGTGAATTGAGAAAGGGTGG	1380
Qy	241	TAACAATTTCACACAGAAACAGGGCGCTGAGAAAAAGCTGAAAGCGCACTGCTTAA	300	Db	1321	ACTGGCCATCACCTCGTGAACCCCGACGTGCGCTACCGTGAATTGAGAAAGGGTGG	1380
Db	241	TAACAATTTCACACAGAAACAGGGCGCTGAGAAAAAGCTGAAAGCGCACTGCTTAA	300	Qy	1381	GCAAACGTGCGCTGCTGTTCTTCACAAAGAACGTCAGTCAGGCTGATGAGCTA	1440
Qy	301	CAATTATCAGACAACTCTGTGGCACTCGACCGAAATTATCGATTAACCTTTATTATA	360	Db	1381	GCAACAGTGCCTGGTTCTTCACAAAGAACGTCAGTCAGGCTGATGAGCTA	1440
Db	301	CAATTATCAGACAACTCTGTGGCACTCGACCGAAATTATCGATTAACCTTTATTATA	360	Qy	1441	CGACAAAGCCTACTACAGTTGGCAGTCAGTCAGTCAGGCTTATTCAAGGGGATT	1500
Qy	361	AAAATTAAAGGGTATATTAATGATCGATTAATAAGGGATAAACCATGGTGA	420	Db	1441	CGACAAAGCCTACTACAGTCAGTCAGTCAGGCTTATTCAAGGGGATT	1500
Db	361	AAAATTAAAGGGTATATTAATGATCGATTAATAAGGGATAAACCATGGTGA	420	Qy	1501	GATTGTAACACCTCGCGGAAATCTAGGGPTGGATCTGAAAGCTGATGATTCA	1560
Qy	421	CAAGGGAGGGAGCTGTACCGGGTGTGACCTGGTACCGGGGACGT	480	Db	1501	GATTGTAACACCTCGCGGAAATCTAGGGPTGGATCTGAAAGCTGATGATTCA	1560
Db	421	CAAGGGAGGGAGCTGTACCGGGTGTGACCTGGTACCGGGGACGT	480	Qy	1561	GTTCTGACTCTGTGAAGGTGAAACCGGGCATCTGGATGCGAAGGAGCTACACTTAGT	1620
Qy	481	AAACGCCAACAGTCAGCGCTGTCGGCGAGGGGATGCGAACCTACGGCAAGGT	540	Db	1561	GTTCTGACTCTGTGAAGGTGAAACCGGGCATCTGGATGCGAAGGAGCTACACTTAGT	1620
Db	481	AAACGCCAACAGTCAGCGCTGTCGGCGAGGGGATGCGAACCTACGGCAAGGT	540	Qy	1621	GATTAAAGAATTGAGCGATAAGGGATCAACAGTTACAGTTAGATAACCGCAAT	1680
Qy	541	GACCTCTGAAGTCATCTGACACCCGAACTCTGCCCTGTGAC	600	Db	1621	GATTAAAGAATTGAGCGATAAGGGATCAACAGTTACAGTTAGATAACCGCAAT	1680
Db	541	GACCTCTGAAGTCATCTGACACCCGAACTCTGCCCTGTGAC	600	Qy	1681	GTGGACACCGCTCTGGCGAAGTAGATGGCAAGATAAGGGATCAACAGTTACAGTTAGATAACCGCAAT	1740
Qy	601	CACCTTGGTACGGCTCGTCAAGTCAGCTGGCGGAGGGCACTACGGCA	660	Db	1681	GTGGACACCGCTCTGGCGAAGTAGATGGCAAGATAAGGGATCAACAGTTACAGTTAGATAACCGCAAT	1740
Db	601	CACCTTGGTACGGCTCGTCAAGTCAGCTGGCGGAGGGCACTACGGCA	660	Qy	1741	CAAATTCGAAGCTGGTTATCGCCAACAGATGCGATGCGAATGGGGGCTTGAAGCGGT	1800
Qy	661	CTTCCTCAAGTCGCCATGCCGAAGGCTAGTCAGTCAGGAGGCCATCTTCAGGA	720	Db	1741	CAAATTCGAAGCTGGTTATCGCCAACAGATGCGATGCGAATGGGGGCTTGAAGCGGT	1800
Db	661	CTTCCTCAAGTCGCCATGCCGAAGGCTAGTCAGTCAGGAGGCCATCTTCAGGA	720	Qy	1801	GAAAGCTACACAAAGTCGCCATTCGGCTGATGGCGCTGCCAGAGCGCT	1860
Qy	721	CGACGGCAACTACAAGACCCGGCCSAGTGAAAGTGCAGGCGACACCTGGTAAACCG	780	Db	1801	GAAAGCTACACAAAGTCGCCATTCGGCTGATGGCGCTGCCAGAGCGCT	1860
Db	721	CGACGGCAACTACAAGACCCGGCCSAGTGAAAGTGCAGGCGACACCTGGTAAACCG	780	Qy	1861	GCCCCGCTGTTGAAATCCGGCTGATGGCGCTGCCAGACCCCTGAAACACGGC	1920
Qy	781	CATCGAGCTGAGGGCATGAGCTTCAAGGAGCGCAACATCTGGGACAAACCTGG	840	Db	1861	GCCCCGCTGTTGAAATCCGGCTGATGGCGCTGCCAGACCCCTGAAACACGGC	1920
Db	781	CATCGAGCTGAGGGCATGAGCTTCAAGGAGCGCAACATCTGGGACAAACCTGG	840	Qy	1921	GAAAGGGACCTTTGAACTGGGAAAACCTGGCCATGGTAAAGGTGGCTGATGGCAC	1980
Qy	841	GTACAACACTACAACAGGCCAACAGTCTTATCATGGCCGACAGGCAATCAA	900	Db	1921	GAAGGGACCTTTGAACTGGGAAAACCTGGCCATGGTAAAGGTGGCTGATGGCAC	1980
Db	841	GTACAACACTACAACAGGCCAACAGTCTTATCATGGCCGACAGGCAATCAA	900	Qy	1981	CAACTCGAAAATCGCAACAAAGTGGCTCCGGTACCTTATGGGGTAGATAAACGCAA	2040
Qy	901	GCTGAACCTCAAGATCGGCCAACACATCGAGGACGCCAGCTGCCGACACTA	960	Db	1981	CAACTCGAAAATCGCAACAAAGTGGCTCCGGTACCTTATGGGGTAGATAAACGCAA	2040
Db	901	GCTGAACCTCAAGATCGGCCAACACATCGAGGACGCCAGCTGCCGACACTA	960	Qy	2041	CCTGGTGAATTCTGGTGAATTGAGTGGCTGATGGTAACTGGCACAAATTTCCTGGTGTGAGA	2160
Qy	961	CGACGAGAACCCCCATCGGCCAGCGCCGAGCTGCCGACACTCTGGTGA	1020	Db	2041	CCTGGTGAATTCTGGTGAATTGAGTGGCTGATGGTAACTGGCACAAATTTCCTGGTGTGAGA	2160
Db	961	CGACGAGAACCCCCATCGGCCAGCGCCGAGCTGCCGACACTCTGGTGA	1020	Qy	2101	CCCAATTCTGGTGAATTGAGTGGCTGATGGTAACTGGCACAAATTTCCTGGTGTGAGA	2160
Qy	1021	CTACCAACTCCGCCCTGAGCAAAGAACCTACACTCTGGTGA	1080	Db	2101	CCCAATTCTGGTGAATTGAGTGGCTGATGGTAACTGGCACAAATTTCCTGGTGTGAGA	2160
Db	1021	CTACCAACTCCGCCCTGAGCAAAGAACCTACACTCTGGTGA	1080	Qy	2161	GGTGAAGGGTGAATGCCAACATAGCGAAAACCTTAAATTGAGTGGCTGAGA	2220
Qy	1081	GTTCGTGAAACCCGGCGGGGATCACTCTGGCATGGAGCTGCTGA	1140	Db	2161	GGTGAAGGGTGAATGCCAACATAGCGAAAACCTTAAATTGAGTGGCTGAGA	2220
Db	1081	GTTCGTGAAACCCGGCGGGGATCACTCTGGCATGGAGCTGCTGA	1140	Qy	2221	ACTACCTGTCCTACCTCCAAACACTGTCACCTTCTCCTGTTGAGTGGCTTC	2280
Qy	1141	TACTCGCATTCGTTGTAACAACTATAAGTAGCTGAACTTGTCTGTAGTGGCAA	1200	Db	2221	ACTACCTGTCCTACCTCCAAACACTGTCACCTTCTCCTGTTGAGTGGCTTC	2280
Db	1141	TACTCGCATTCGTTGTAACAACTATAAGTAGCTGAACTTGTCTGTAGTGGCAA	1200	Qy	2281	CGCTPATCCGGATCATATGAAACGGCATGACTTCAGAGTGGCATGCCAGGTTA	2340

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 2, 2005, 18:19:33 ; Search time 201.5 Seconds
(without alignments)

6426.229 Million cell updates/sec

Title: US-10-649-433C-1

Perfect score: 1920

Sequence: 1 gtttgcacgttatcatcgat.....agttagcggaaattgatctg 6729

Scoring table: BLOSSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
DelOp	6.0	Delext	7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p_model1 -DEV=xIP
-Q=/con2.1/lsp2o_spool/p/US10649433/runat_02082005_164945_1532/app_query.fasta_1.6919
-DB=PIR -QFMT=fasta -SUFFIX=rpx -MINMATCH=0.1 -LOCPEXT=0 -LOCPECL=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40_cdi -LIST=45
-DOCAALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLN=0 -MAXLN=2000000000
-USER=US10649433 @CGN 1.1 332 @RUN 2002005_164945_1532 -ICPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT=0 -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : PIR_79:
1: pir1:
2: pir2:
3: pir3:
4: pir4:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1765	14.8	360	1	RPECL	lactose operon rep
2	1765	14.8	360	4	S58116	lac repressor - sy
3	1754	14.7	360	2	P90678	lac operon transcr
4	1747	14.7	360	2	B88529	lac operon transcr
5	1579	13.2	332	1	JGBIG	D-galactose-bindin
6	1579	13.2	332	1	B91009	galactose-bindin
7	1579	13.2	332	2	D88853	D-galactose-bindin
8	1537	12.9	332	1	S15554	D-galactose-bindin
9	1510	12.7	332	2	AF0781	galactose-bindin
10	1508	12.7	332	1	S20390	beta-lactamase (EC)
11	1467	12.3	286	2	T31301	beta-lactamase (EC)
12	1467	12.3	286	2	S47061	beta-lactamase (EC)
13	1467	12.3	286	4	S41975	beta-lactamase (EC)
14	1464	12.3	286	1	PNECP	beta-lactamase (EC)

RESULTS

RESULT 1

RPECL

lactose operon repressor - Escherichia coli (strain K-12)

N;Alternate names: lac repressor

C;Species: Escherichia coli

C;Date: 30-Apr-1982 #Sequence_revision 30-Apr-1982

C;Accession: A93198; S40651; A92122; A93785; JC2566

R;Parabagh, P. J.

Nature 274, 765-769, 1978

A;Title: Sequence of the lacI gene.

A;Reference number: A93198; MUID:78246991; PMID:355891

A;Accession: A93198

A;Molecule type: DNA

A;Residues: 1-360 <P>

A;Cross-references: UNIPROT:P03023

R;Beyreuther, K.; Adler, K.; Fanning, E.; Murray, C.; Geissler, N.

Eur. J. Biochem. 59, 491-509, 1975

A;Title: Amino-acid sequence of lac repressor from Escherichia coli. Isolation, sequencing and comparison with the lac operon repressor

A;Reference number: A91234; MUID:76091932; PMID:1107032

A;Accession: A91234

A;Molecule type: protein

A;Residues: 1-147;159-163; 'Q' 165-230;233-360

A;Note: the active repressor is a tetramer of identical chains

A;Note: this protein was obtained from a strain with the I-SQ mutation, which leads to a R;Beyreuther, K.

Nature 274, 767, 1978

A;Contents: annotation; revision

A;Note: this sequence has since been revised

R;Beyreuther, K.

Eur. J. Biochem. 59, 491-509, 1975

A;Title: Amino-acid sequence of lac repressor from Escherichia coli. Isolation, sequencing and comparison with the lac operon repressor

A;Reference number: A91234; MUID:76091932; PMID:1107032

A;Accession: A91234

A;Molecule type: protein

A;Residues: 1-147;159-163; 'Q' 165-230;233-360

A;Note: the active repressor is a tetramer of identical chains

A;Note: this protein was obtained from a strain with the I-SQ mutation, which leads to a R;Beyreuther, K.

Nature 274, 767, 1978

A;Contents: annotation; revision

A;Note: this sequence has since been revised

R;Beyreuther, K.

J. Biol. Chem. 248, 110-121, 1973

A;Reference number: A92122; MUID:73143730; PMID:4571224

A;Accession: A92122

A;Molecule type: protein

A;Residues: 1-59;96-101;206-215;328-347 <PLA>

A;Note: removal of residues 1-59 and 328-347 results in the molecule losing its DNA-bindin

R;Ganem, D.; Miller, J.H.; Files, J.G.; Platt, T.; Weber, K.

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 5, 2005, 01:59:21 ; Search time 19523 Seconds
(without alignments)

13119.627 Million cell updates/sec

Title: US-10-649-433C-1

Perfect score: 6729

Sequence: 1 gtttacagcttatcatgatcg.....agttagccgttgcattatgt 6729

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 190321347000 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Result No.	Score	Query	Match	Length	DB	ID	Description
C 1	1055.8		15.7	1070	1	AJ281552	AJ281552 4A3A-P6F1
C 2	1048.2		15.1	1048	7	C0552396	ACLY 50
C 3	984.8		14.6	1013	4	BM438846	IPLvrc0015
C 4	924.2		13.7	1004	1	AJ281480	4A3A-P4G8
C 5	889.8		13.2	928	7	CO487414	GQ0227_B7
C 6	845.8		12.6	902	7	CR753463	DKF2p169G
C 7	837.8		12.5	841	1	AL042026	DKP2p334E
C 8	835.8		12.4	935	4	BG838279	Gc01_10e0
C 9	823.2		12.2	854	4	BH438950	IPLvF0049
C 10	820.2		12.2	865	7	CK18014	21.8Bn06_P1
C 11	812.8		12.1	1126	8	B2572566	msh2_15.1
C 12	806.4		12.0	819	6	CD649375	CvGnd0008
C 13	800.6		11.9	856	7	CN823189	Oa_sp1bn
C 14	796		11.8	1011	8	B2576726	msh2_5071
C 15	791.4		11.8	827	7	CR823902	Oa_sp1bn
C 16	785.4		11.7	1574	8	B2572566	msh2_2693
C 17	779.4		11.6	1073	7	CF26952	Fcy1c0_id8
C 18	778		11.6	794	5	BQ751936	EST632499
C 19	777.8		11.6	800	1	AJ281449	4A3A-P4D5
C 20	773.2		11.5	1336	8	B2575810	msh2_4637
C 21	772.8		11.5	789	6	CD280320	G44224_42
C 22	772		11.5	846	7	CV468077	est_1_van
C 23	770.6		11.5	779	7	CK394523	hggad4D0
C 24	770.4		11.4	1249	8	B2572284	msh2_2572

Query Match	Score	Length	DB	EST
Best Local Matches 1068; Conservative 0;	Score 1055.8; Pred. No. 1.2e-272;	Length 1070;	Db	EST 01-SEP-2004.
Matches 1068; Conservat 0;	Mismatches 2;	Indels 1;	Gaps 1;	mRNA sequence.
Qy 3959 GTTGGCGAACTTAACTAAGTGGAAACTTACCTGTTCCGGAAACAACTTAATAGA 4018	Db 1070 GTTGGCGAACTTAACTAAGTGGAACTTACCTGTTCCGGAAACAACTTAATAGA 1011	Db 4019 CTGGATGGGGATAAAGTTCAGGACCACTCTGGCTCCGGCTTGCTG 4078	Db 1010 CTGGATGGGGATAAAGTTCAGGACCACTCTGGCTGG-CCTTCCGGCTGCTG 952	RESULT 2 COS52396.c LOCUS C0552396_c DEFINITION AcLy4_50 Sea lamprey AcLy Petromyzon marinus cDNA, VERSION C0552396 KEYWORDS EST. SOURCE Petromyzon marinus (sea lamprey) ORGANISM Petromyzon marinus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia; Petromyzontiformes; Petromyzontidae; Petromyzon.
Qy 4079 GTTATTGCTGATAAATCTGGGCCGTAGCTGGGGTCTCCGGTATCATTCAGACGGGAAC 4138	Db 951 GTTATTGCTGATAAATCTGGGCCGTAGCTGGGGTCTCCGGTATCATTCAGACGGGAAC 892	Db 4139 GGGGCCAGATGATAAGGCCCTCCCGTATCGTAGTGTATCATTCAGACGGGAAC 4198	Db 891 GGGGCCAGATGATAAGGCCCTCCCGTATCGTAGTGTATCATTCAGACGGGAAC 832	REFERENCE 1 AUTHORS Pancer, Z.; Mayer, W.E.; Klein, J. and Cooper, M.D. TITLE prototypic T-cell receptor and CD4-like coreceptor expressed in lymphocytes of the agnathan sea lamprey JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101, 13273-13278 (2004) COMMENT Contact: Pancer, Zeev Division of Developmental and Clinical Immunology The University of Alabama at Birmingham 378 Wallace Tumor Institute, 1530 Third Avenue, South, Birmingham, AL 3524-3300 Tel: 205-975-5812 Fax: 205-975-7218 Email: zpancer@uab.edu.
Db 4199 TATGGATGAACTAAATAGACAGATCGTGTAGATGGTGCCTACTGATTAAGCATTTGTA 4258	Db 831 TATGGATGAACTAAATAGACAGATCGTGTAGATGGTGCCTACTGATTAAGCATTTGTA 772	Db 4259 ACTGTCAAGCCAAGTAACTCATATATACCTTATAGTGTAAACTCTCATTTTTAAT 4318	Db 771 ACTGTCAAGCCAAGTAACTCATATATACCTTATAGTGTAAACTCTCATTTTTAAT 712	FEATURES 1. source /organism="Petromyzon marinus" /mol_type="mRNA" /db_xref="taxon:7157" /cell_type="lymphocyte" /dev_stage="immune stimulated larvae" /note="vector: pGEM-T Easy; lymphocyte mRNA ESTs from PCR subtracted cDNA libraries of immune stimulated larvae. All are single pass 5' or 3' sequences randomly cloned in pGEM-T Easy (Promega)." /Location/Qualifiers
Qy 4319 TAAAAGGATCTGGTGAAGATCCTTTGATAATTCATGACAAATTCCTTAACCTGA 4378	Db 711 TAAAAGGATCTGGTGAAGATCCTTTGATAATTCATGACAAATTCCTTAACCTGA 652	Db 4379 GTTTTCCTTCACTGACGGTCAAGCCGGTACAAGGAAATCAAGGATCTCTGTGAGATC 4438	Db 651 GTTTTCGTTCCACTGACGGTCAAGCCGGTACAAGGAAATCAAGGATCTCTGTGAGATC 592	ORIGIN Query Match 15.1%; Score 1018.2; DB 7; Length 1048; Best Local Similarity 99.2%; Pred. No. 1.6e-262; Matches 1043; Conservative 1; Mismatches 4; Indels 8 3; Gaps 2
Qy 4439 TTTCCTCTGGCGTAATCTGTGCTGTCACACAAAAACCCGGTACACGGCTGGTGT 4498	Db 591 TTTCCTCTGGCGTAATCTGTGCTGTCACACAAAAACCCGGTACACGGCTGGTGT 532	Db 4499 TTGTTCGCGGATCAAGAGCTTACCAACTCTTTCGGAAAGTAAACGGCTTCAAGCAGC 4558	Db 531 TTGTTCGCGGATCAAGAGCTTACCAACTCTTTCGGAAAGTAACTGGCTTCAAGCAGC 472	Db 4559 GCAGATAACAAATACCTGTCCTCTGTGCTGTAAGCCACATTCAAGAACCTC 4618
Db 471 GCAGATAACAAATACCTGTCCTCTGTGCTGTAAGCCACATTCAAGAACCTCAGAACCTC 412	Db 4619 TGTAGCACCGCCCTACATACCTGTCCTCTGTGCTGTAAGCCACATTCAAGAACCTC 4678	Db 471 TGTAGCACCGCCCTACATACCTGTCCTCTGTGCTGTAAGCCACATTCAAGAACCTC 352	Db 479 CGATAAGTGTGCTCTGCTTACCGGGTTGGACTCAAGAGTAGTAACTGAGCTTCAAGCAGC 4738	Db 351 CGATAAGTGTGCTCTGCTTACCGGGTTGGACTCAAGAGTAGTAACTGAGCTTCAAGCAGC 292
Qy 4739 GTGGGGTGAAGGGGGTCTGGTGTAGTCTGTGACACGCCAGTGGCAAGACCTAACCGA 4798	Db 411 TGTAGCACCGCCCTACATACCTGTCCTCTGTGCTGTAAGCCACATTCAAGAACCTC 352	Db 479 CGACAGCTTACGGTAAGGCCAGGGTCTGGTGTAGTCTGTGACACGCCAGTGGCAAGACCTAACCGA 232	Db 351 CGATAAGTGTGCTCTGCTTACCGGGTTGGACTCAAGAGTAGTAACTGAGCTTCAAGCAGC 292	Db 4223 CGCTGAGATAGGGTCTCTGCTTACGTGATTAAGCTGAGCTTCAAGCAGCAAGTTTACTCTATA 4281
Qy 4799 ACTGAGATACTTACAGGGTAGGCTATGAAAGAGGCCAGCTTCCGGAAAGGAAAGGC 4858	Db 231 ACTGAGATACTTACAGGGTAGGCTATGAAAGAGGCCAGCTTCCGGAAAGGAAAGGC 172	Db 4799 ACTGAGATACTTACAGGGTAGGCTATGAAAGAGGCCAGCTTCCGGAAAGGAAAGGC 4918	Db 868 CGCTGAGATAGGGTCTCTGCTTACGTGATTAAGCTGAGCTTCAAGCAGCAAGTTTACTCTATA 809	Db 4283 TATAGCTTGTAGTGTAACTTAAAGGACTCTAGGTGAAAGATCT 434
Qy 4859 GGACAGCTTACGGTAAGGCCAGGGTCTGGTGTAGTCTGTGACACGCCAGTGGCAAGACCTAACCGA 5029	Db 171 GGACAGCTTACGGTAAGGCCAGGGTCTGGTGTAGTCTGTGACACGCCAGTGGCAAGACCTAACCGA 112	Db 4799 ACTGAGATACTTACAGGGTAGGCTATGAAAGAGGCCAGCTTCCGGAAAGGAAAGGC 4978	Db 748 TTTCTGATAATCTGACCAAAATCCCTTAACGTGTTCTGCTGAGCTTCAAGCAGCAAGTTTACTCTATA 689	Db 4403 CCCGGTAGAAAGATCAAAGGATCTCTGAGATCTTCTGAGCTTCACTGAGCTTCACTGCTG 4462
Qy 4919 GGAAAAGCCCTGCTTACGGGAGGCTTCCAGGGAGGAGCTTCCAGGAGCTTCCAGG 5029	Db 111 GGAAAAGCCCTGCTTACGGGAGGCTTCCAGGGAGGAGCTTCCAGGAGCTTCCAGG 52	Db 4979 ATTTCCTGTAATCTGCTGTCAGGGGGCGGAGGCCCTATGGAAAAGGCCAGCAA 5029	Db 688 CCCGGTAGAAAGATCAAAGGATCTCTGAGATCTTCTGAGCTTCACTGAGCTTCACTGCTG 629	Db 4403 CCCGGTAGAAAGATCAAAGGATCTCTGAGATCTTCTGAGCTTCACTGAGCTTCACTGCTG 4462

Copyright GenCore version 5.1.6
(c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 2, 2005, 18:11:07 ; Search time 1027.5 Seconds
(without alignments)

6707.116 Million cell updates/sec

Title: US-10-649-433C-1

Perfect score: 11920

Sequence: 1 gtttgcagcttatcatcga.....agttagcgccaaattgatctg 6729

Scoring table: BLOSSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_plus_n2p.model -DEV=x1P
-Q=/con2.1/USPTO_spool_P/US10649433/runat_02082005_164945_1521/app_query.fasta_1.6919
-DB=uniProt -QFMT=fasban -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40_cdi -LIST=45
-DOCAALIGN=0.00 -PTHR=SCORE=oct -THR=MAX=100 -THR=MIN=0 -ALIGNC=15 -MODE=LOCAL
-OUTFMT=pfo -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO_MMAP -LARGEQUERY -NEG_SCORE=0.0 -WAIT -DSTPLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=0.5 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-PGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : UniProt_03:
1: uniprot_sprefc:
2: uniprot_trembl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1675	14.8	360	1	LACI_ECOLI	P03023	escherichia
2	1754	14.7	360	2	Q7AH57	O7ah57	escherichia
3	1747	14.7	360	2	Q8X684	Q8x684	escherichia
4	1579	13.2	332	1	DGAL_ECOLI	P02927	escherichia
5	1579	13.2	332	2	Q8X671	Q8x641	escherichia
6	1537	12.9	332	1	DGAL_CITR	P23925	citrobacter
7	1518	12.7	332	1	DGAL_SALTY	P23905	salmonella
8	1510	12.7	332	2	Q8Z5A4	O8z5a4	salmonella
9	1467	12.3	286	2	Q3B058	Q3B058	bacteriophila
10	1467	12.3	286	2	Q00626	Q00626	staphylococ
11	1467	12.3	286	2	Q79CL6	Q79cl6	methylbaci
12	1467	12.3	286	2	Q79DR3	Q79dr3	escherichia
13	1464	12.3	286	1	BLAT_ECOLI	P62593	escherichia
14	1464	12.3	286	1	BLAT_SALTI	P62594	salmonella
15	1464	12.3	286	2	Q6A253	Q6a253	haemophilus
16	1464	12.3	286	2	Q6LB99	Q6lbn9	pseudomonas

ALIGNMENTS

AC	ID	LACI_ECOLI	STANDARD	PRT	360 AA.
P03023	009196	P71309;	Q47338;		
DT	21-JUL-1986	(Rel. 0);	Created		
DT	10-OCT-2003	(Rel. 42);	Last sequence update		
DT	25-JUN-2005	(Rel. 46);	Last annotation update		
DE			Lactose operon repressor.		
GN		Name=laci;	OrderedLocusNames=b03455;		
OS		Escherichia coli			
OC		Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.			
OX	[1]	NCBI_TaxID=562;			
RN	[1]	SEQUENCE FROM N.A.			
RN	[2]	SEQUENCE FROM N.A.			
RP		SEQUENCE FROM N.A.			
RP		STRAIN=K12 / MG1655;			
RP		MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;			
RA		Blattner P.R.; Plunkett G. III, Perna N.T., Burland V.,			
RA		Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA		Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA		Namach A., Telfer P., Roberts D., Schramm S., Davis R.W.,			
RA		Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]	SEQUENCE FROM N.A.			
RP		Strain=K12 / MG1655;			
RC		Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,			
RC		Federer N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,			
RA		Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA		Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA		Namach A., Telfer P., Roberts D., Schramm S., Davis R.W.,			
RA		Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.			
RN	[5]	SEQUENCE FROM N.A.			
RP		Strain=K12 / MG1655;			
RA		Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,			
RA		Federer N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,			
RA		Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA		Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA		Namach A., Telfer P., Roberts D., Schramm S., Davis R.W.,			
RA		Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.			
RN	[6]	SEQUENCE FROM N.A.			
RP		Strain=K12 / MG1655;			
RA		Chen J., Matthews K.K.S.M.,			
RA		Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.			

[5] SEQUENCE FROM N.A.
RP Marsh S.;
RA Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.

[6] SEQUENCE OF 1-147; 159-230 AND 233-360.
RX MEDLINE=73143730; PubMed=1107032;

RA Beyreuther K., Adler K., Fanning E., Murray C., Klemm A., Geisler N.;
RT "Amino-acid sequence of lac repressor from *Escherichia coli*.
RT Isolation, sequence analysis and sequence assembly of tryptic peptides
RT and cyanogen-bromide fragments.";
RL Eur. J. Biochem. 59:491-509(1975).

[7] SEQUENCE OF 1-59; 96-101; 206-215 AND 328-347.
RX MEDLINE=73143730; PubMed=1571224;

RA Platt T., Files J.G., Weber K.;
RT "Lac repressor. Specific proteolytic destruction of the NH₂-terminal
region and loss of the deoxyribonucleic acid-binding activity.";
RL J. Biol. Chem. 248:110-121(1973).

[8] SEQUENCE OF 60-70; 73-78 AND 83-86.
RX MEDLINE=74126378; PubMed=594437;

RA Ganem D., Miller J.H., Files J.G., Platt T., Weber K.; Allen F.L.,
RT "Reinitiation of a lac repressor fragment at a codon other than
AUG";
RT Proc. Natl. Acad. Sci. U.S.A. 70:3165-3169(1973).

[9] SEQUENCE OF 1-60 FROM N.A.
RX MEDLINE=88230449; PubMed=3286877;

RA Gordon A.J.E., Burns P.A., Fix D.F., Yatagai F., Allen F.L.,
RA Horstall M.J., Gray J.A., Bernelet-Moens C.,
RA Glickman B.W.;
RT "Missense mutation in the lacI gene of *Escherichia coli*. Inferences on
the structure of the repressor protein.";
RL J. Mol. Biol. 200:239-251(1988).

RP MUTAGENESIS.
RX MEDLINE=90183956; PubMed=2178920;
RA Lehming N., Sartorius J., Kisters-Woike B., von Wilcken-Bergmann B.,
RA Mueller-Hill B.;
RT "Mutant lac repressors with new specificities hint at rules for
protein-DNA recognition.";
RL EMBO J. 9:615-621(1990).

[11] MUTAGENESIS.
RP MEDLINE=93322386; PubMed=046748;

RA Markiewicz P., Kleina L.G., Cruz C., Ehret S., Miller J.H.;
RT "Genetic studies of the lac repressor. XIV. Analysis of 4000 altered
Escherichia coli lac repressor reveals essential and non-essential
residues, as well as 'spacers' which do not require a specific
sequence.";
RT J. Mol. Biol. 240:421-433(1994).

[12] 3D-STRUCTURE MODELING.
RX MEDLINE=91249837; PubMed=2040302;

RA Kisters-Woike B., Lehming N., Sartorius J., von Wilcken-Bergmann B.,
RA Mueller-Hill B.;
RT "A model of the lac repressor-operator complex based on physical and
Genetic data.";
RL Eur. J. Biochem. 198:411-419(1991).

[13] STRUCTURE BY NMR.
RX MEDLINE=89113344; PubMed=3064080;

RA Boelens R., Lameriche R.M.J.N., Rullmann J.A.C., van Boom J.H.,
RA Kaptein R.;
RT "The interaction of lac repressor headpiece with its operator: an NMR
view.";
RL Protein Seq. Data Anal. 1:487-498(1998).

[14] STRUCTURE BY NMR.
RP MEDLINE=89108867; PubMed=2742823;

RA Lameriche R.M.J.N., Boelens R., van der Marel G.A., van Boom J.H.,
RA Kaptein R., Buck F., Fera B., Rueterjans H.;